Jinhui Wang • Bernard Ricardo

Competitive Physics

Mechanics and Waves

by
Physics Olympiad medalists and trainers

World Scientific

Competitive Physics

Mechanics and Waves

This page intentionally left blank

Competitive Physics

Mechanics and Waves

Jinhui Wang

Bernard Ricardo

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Wang, Jinhui (Former Physics Olympiad student), author. | Ricardo, Bernard, 1985– author.
Title: Competitive physics : mechanics and waves / Jinhui Wang (Hwa Chong Junior College, Singapore), Bernard Ricardo (NUS High School of Mathematics and Science, Singapore).
Description: Singapore ; Hackensack, NJ : World Scientific, [2018]
Identifiers: LCCN 2017048390| ISBN 9789813233942 (hardcover ; alk. paper) | ISBN 981323394X (hardcover ; alk. paper)
Subjects: LCSH: Physics--Problems, exercises, etc. | Physics--Competitions.
Classification: LCC QC32 .W216 2018 | DDC 530/.076--dc23
LC record available at https://lccn.loc.gov/2017048390

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/10817#t=suppl

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

Dedication

The Physics Olympiads were some of my most enjoyable experiences in high school and I hope to share the elegance of problem-solving, prevalent in such competitions, with more people. Writing this book has really been a fulfilling journey that has evoked many moments of nostalgia, ranging from the thrill of solving a problem to my past struggles when learning about physics. I am extremely grateful to the following people who have supported me in the course of writing — during my hardest times in National Service — in one way or another.

My co-author, Bernard Ricardo. I can safely say that he is one of the best physics educators in Singapore. His enthusiasm and illuminating pedagogy during his lessons for the Physics Olympiad national training team inspired me to pursue and teach physics. It has been a joy working with him and his many insights have added a novel perspective to this book.

My teachers, Tan Jing Long, Mr Kwek Wei Hong and Mrs Ng Siew Hoon. I still remember the quirky jokes that the class makes during your lessons. Thank you for introducing me to physics and infecting me with your love for the subject.

My students, Cui Zizai, Mao Ziming, Chang Hexiang, Guo Yulong, Andrew Ke Yanzhe, Ryan Wong Jun Hui and many more. It was a pleasure to teach you and your inquisitiveness in our lessons together was a great source of motivation during my conscription. I am also grateful to you for proofreading chapters of this book and providing constructive feedback.

My friends, Jiang Yue and Ma Weijia. Thank you for taking time out from your busy university schedules to draw numerous diagrams for this book as I was preoccupied with National Service. This milestone would literally not have been possible without you.

My parents. Thank you for everything that you have done for me, from cooking meals at home to discussing life decisions with me. I would not be who I am today without you.

Finally, it is with the greatest joy that I present *Competitive Physics* and I hope that you enjoy the book!

— Wang Jinhui

It is a pleasure for me to present our book, *Competitive Physics*. The process of this book's production, that has been thrilling and satisfying, could not be completed without these special people in my life whom I would like to extend my deepest, most heart felt thanks.

My co-author, Wang Jinhui. This book was birthed out of his vision. He is one of the brightest young men that I know and it was extremely enjoyable to put our thoughts together. I hope this would be the start of many more collaborations and substantial discussions.

My mentor, Professor Yohanes Surya Ph.D. I was first introduced to the world of Physics Olympiads by him. Full of passion, he dedicated himself to impart his knowledge and love for physics to so many people. His slogan, *Physics is Fun*, has encouraged and inspired me to impact the world of Physics Education.

My wife, Yoanna Ricardo, daughter, Evangeline Ricardo, and son, Reinard Ricardo. Thank you so much for always supporting and praying for the completion of this book. The times at home and on the road with you are the times when I was most fruitful in writing and I am truly grateful for those good memories.

For all our readers, I hope you will find this book enjoyable and have fun reading it as we have had fun writing it!

— Bernard Ricardo

Preface

Competitive Physics grew out of a Physics Olympiad course taught by Wang Jinhui at Hwa Chong Institution — intended to prepare students for the annual Physics Olympiads and to imbue deeper knowledge in physics beyond the typical high school syllabus. It quickly became a collaboration with his former trainer in the Singapore Physics Olympiad national training team, Bernard Ricardo.

Competitive Physics is meant to be a theory-cum-problem book. The first half of each chapter explores physical theories with illustrations of how they can be creatively applied to problems. The latter half of each chapter revolves around puzzles that we hope will intrigue readers, as we believe that problem-solving is a crucial process in grasping the subtleties of the contents. Therefore, we have included a multitude of problems which are ranked by increasing difficulty from one to four stars. Some problems are original; some are taken from the various Physics Olympiads while the others are instructive classics that have withstood the test of time.

This book is the first part of a two-volume series which will discuss general problem-solving methods and delve into mechanics and waves — setting a firm foundation for other topics that will be presented in the second volume.

We envision problem-solving to be a fun process — from the initial excitement of approaching an unfamiliar problem, to the joy of pitting all of one's knowledge against it and finally, the satisfaction earned from solving it after numerous failed attempts. In light of this, our goal is to spread the passion of problem-solving — an infectious hobby. It is difficult to quantify the factors that make a problem interesting or elegant but the following have been our guiding principles in writing *Competitive Physics*:

1. Physical Significance. Quintessentially, physics is about modeling the world around us. Therefore, it is gratifying to be able to analyze everyday phenomena and to leverage on this knowledge to improve such processes.

For example, a problem in Chapter 3 asks: how should we run towards a shelter to minimize how wet we get in the rain? Meanwhile, Chapter 6 elucidates the reason behind why two colliding billiard balls leave at right angles relative to each other.

2. Intuition. There are many overarching themes in physics — symmetry, the equivalence of different observational frames of reference, reversibility of processes and many more. Not only are these useful as sleights-of-hand in problem-solving, they reveal crucial aspects of the common structure of physical theories. Developing a strong hunch for them — a gut feeling that constantly bugs you to search for ways to exploit them — may prove to be beneficial in one's future physics journey. As such, we have devoted the entire first chapter, Minimalistic Arguments, to honing this physical intuition.
3. Insight. Sometimes, a seemingly complex problem can be vastly simplified by making an astute observation — whether mathematical or physical. Perhaps, it is to express the solution in terms of vectors or perhaps it is to observe that two different scenarios "feel" the same to a certain entity and thus conclude that the entity will respond in the same manner in both cases. Maybe it is to draw enlightening analogies between two problems that appear to be completely disparate on the surface. Ultimately, such problems which require perceptive thought do not have cookie-cutter approaches and require the reader to invent an appropriate technique on the spot. They hence implore the reader to really think and are very rewarding to solve.
4. Fundamentals. The objectives above would not be possible without first mastering the fundamentals of a theory — the situations that it can be validly applied to, its assumptions and its ramifications. As such, we have also included many classic problems to reinforce understanding of the basics. To this end, we are extremely grateful to Dr. David J. Morin for allowing us to use some problems from his exemplary textbook: *Introduction to Classical Mechanics.*

In summary, our guiding principles are "PIIF", as in the onomatopoeia "pffft" when, having read this book, you scoff at a future problem after swiftly spotting its trick. Jokes aside, it is paramount for the reader to first attempt the problems before peeking at the solutions. Even when perusing the solution to a problem, the reader should inspect it line by line until he or she reaches an inspiration that sets him or her back on track in attempting the problem again. Only by experiencing the process of problem-solving yourself can you internalize the clues in a problem that hint at a certain approach, understand why certain approaches are incorrect or desirable and ultimately, improve. There is no short-cut to developing an intuition for problem-solving besides trudging through an arduous but fulfilling journey of enigmas.

Despite our best efforts; the probability of this book being error-free is, unfortunately, akin to the odds of observing a car plate that reads "PHY51C." Therefore, if the reader does spot any mistakes or dubious points in our discussions, we would appreciate if they are highlighted to us via the email competitivephysicsguide@gmail.com.

This page intentionally left blank

Contents

Chapter 1

Minimalistic Arguments

This chapter will explore the ramifications of a few elementary physical principles to develop the necessary intuition in approaching many physics problems.

1.1 Dimensional Analysis

The units of a quantity essentially encapsulate what it describes. For example, if a quantity has units in joules (J), we can surmise that it is referring to some form of energy. Conversely, if we already know the physical meaning of the quantity and hence know its units, we can guess a solution for the quantity in terms of other given parameters — solely by inspecting their units. This is because the two sides of an equation must first be dimensionally homogeneous (i.e. have the same units) in order for the equality to be valid. Furthermore, operations such as addition and subtraction can only be performed on commensurable quantities (quantities of the same dimensions). For instance, it makes no sense to say whether 1 second is equal to 1 meter or to add 1 Celsius to 1 Coulomb. Lastly, certain operations such as trigonometric functions, exponentials and logarithms can only be performed on dimensionless variables. In any case, there are certain restrictions, pertaining to the units of variables, imposed on every equation that need to be fulfilled to procure physical meaning and ensure logical coherency.

The following is a concrete example of applying dimensional analysis to a problem. If asked to deduce the rest energy of a particle when given its mass M and the speed of light in vacuum c, one can first observe that the dimensions of energy are $\frac{[M][L]^2}{[t]^2}$ which can be read as some mass unit multiplied by some length unit squared and divided by some time unit squared. Thus, one can guess a solution of the form

$$E_{rest} = Mc^2,$$

as the right-hand side also has the same dimensions. This is in fact the correct answer (though we may not even understand what rest energy means, other than the fact that it refers to some form of energy)! Usually, it is wise to express every variable in terms of a common basis comprising convenient dimensions denoted by square brackets before applying dimensional analysis. A common guess for a quantity Q, given parameters x_1, $x_2, \ldots, x_n$, would be of the form

$$Q = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n},$$

where the α's are constants. This is also known as the power law. Then, a set of linear equations can be derived to solve for the α's. For example, the angular frequency, ω of a simple pendulum undergoing small-angle oscillations can be deduced in terms of its physical variables that include the mass of the bob m, length of the string l and the gravitational field strength g. Presuming that ω is independent of the initial angular displacement, we can guess a solution of the following form:

$$\omega = m^\alpha l^\beta g^\gamma,$$

where α, β and γ are constants. We observe that the desired variable and the physical parameters have dimensions of

$$\begin{aligned} \omega &= [t]^{-1}, \\ m &= [M], \\ l &= [L], \\ g &= [L][t]^{-2}. \end{aligned}$$

Substituting these units into the power law expression for ω would yield

$$[t]^{-1} = [M]^\alpha \cdot [L]^\beta \cdot [L]^\gamma [t]^{-2\gamma} = [M]^\alpha [L]^{\beta+\gamma} [t]^{-2\gamma}.$$

Comparing the exponents,

$$\begin{aligned} \alpha &= 0, \\ \beta + \gamma &= 0, \\ -2\gamma &= -1. \end{aligned}$$

Solving,

$$\begin{aligned} \alpha &= 0, \\ \beta &= -\frac{1}{2}, \\ \gamma &= \frac{1}{2}. \end{aligned}$$

Thus, we surmise that

$$\omega = \sqrt{\frac{g}{l}},$$

which happens to be the correct solution! Perhaps, a rather cumbersome process in dimensional analysis is the conversion of the units of all parameters into a standardized set of units. For example, one might wish to convert Joules (J) into SI units. In such cases, considering physical laws can often be very helpful. Suppose we wish to determine the expression of Joules in SI units, we can make use of the work-energy theorem which provides us with the information that the units of force multiplied by that of displacement give Joules. Furthermore from $F = ma$, we conclude that the units of force (a Newton) is $kgms^{-2}$ in SI units. Thus, Joules can be expressed as kgm^2s^{-2} in terms of SI units or dimensions $[M][L]^2[t]^{-2}$.

1.1.1 *Limitations*

Obviously, dimensional analysis is not omnipotent and cannot be applied to solve all problems accurately. Well, if this were not the case, we will only be examining the units of variables in every question and do not need any knowledge of physics. We shall elaborate on the limitations of dimensional analysis in this section.

Firstly, dimensional analysis will not provide us with any information about dimensionless parameters, such as constants, trigonometric functions and exponents. In the previous example, we could have guessed $\omega = 2\pi\sqrt{\frac{g}{l}}$ and it would still have units in s^{-1}. When asked to deduce the formula for the kinetic energy of an object in terms of its mass M and velocity v, one could possibly hypothesize a solution of the form K.E. $= Mv^2$ which is obviously wrong. In light of this fact, we usually include a dimensionless constant k in front of our power law solution.

Similar to the previous point, if a combination of the physical parameters in the form of a power law is dimensionless, there is no unique solution to the set of linear equations. For example, the capacitance of a parallel plate capacitor is $\varepsilon_0 \frac{A}{d}$ where A is the area of the plates and d is the separation between them but it could also be $\varepsilon_0 \sqrt{A}$ and even $\varepsilon_0 \frac{A^{\pi+1}}{d^{2\pi+1}}$ if based purely on dimensional analysis! This is because $\frac{A}{d^2}$ is a dimensionless quantity and we can multiply any power of this to anything in an expression and it would not change the overall units. To account for this, one usually multiplies the power law by an additional factor $f(\frac{A}{d^2})$ that is an arbitrary function of the combinations of parameters that are dimensionless ($\frac{A}{d^2}$ in this case).

An extreme example where dimensional analysis yields a ludicrous result, if we miss out this factor, would be in guessing the solution to the instantaneous displacement of a simple harmonic motion as $x = x_0$, where x_0 is the initial displacement, and missing out the harmonic part of the solution completely!

Perhaps, the biggest limitation of dimensional analysis is its inability to aid in the development of new theories. Suppose that the ideal gas law has not been discovered yet and we wish to find a relationship between the pressure, volume, number of moles and temperature of a gas which are denoted by p, V, n and T respectively. It is impossible to even determine a non-trivial and homogeneous equality relating the variables in this case via dimensional analysis alone. No combination of these parameters can yield a meaningful and valid equation. However, suppose that we already know that $p \propto \frac{nT}{V}$, we can conduct physical experiments to determine whether the "constant" of proportionality is really a physical constant and find its value. Then, we can name this physical constant R and state the ideal gas law as $pV = nRT$. Essentially, dimensional analysis fails in this case as there is a need to introduce a new physical constant with non-trivial dimensions in order to relate the parameters, while dimensional analysis already assumes that the parameters can already be interconnected in a meaningful way.

Finally, it is important to note that a major pitfall lurks when the quantity that we wish to analyze blows up to infinity. For example, suppose that we wish to compute the total distance l_{max} covered by a particle, of mass m and initial velocity v_0, undergoing one-dimensional motion under a quadratic drag force of the form $F_{drag} = kv^2$ where k is the drag constant and v is the particle's instantaneous velocity. The possible parameters are $m = [M]$, $v_0 = [L][t]^{-1}$ and $k = \frac{F}{v^2} = \frac{[M][L][t]^{-2}}{[L]^2[t]^{-2}} = [M][L]^{-1}$. A natural hypothesis based on dimensional analysis for $l_{max} = [L]$ would be

$$l_{max} = \frac{m}{k}$$

but we know that this must be wrong as l_{max} does not scale with the initial velocity v_0. We would expect that l_{max} of a particle with initial velocity $v_0' > v_0$ would be larger than that with initial velocity v_0 as the particle would have travelled some distance during the time it takes to decelerate to velocity v_0, after which it will traverse the same distance as the particle which started with v_0. The reason behind this absurd answer is that there is in fact no upper bound on l_{max}! The particle travels to infinity. Dimensional analysis becomes useless when quantities blow up as they cannot be meaningfully related to given parameters, which are finite most of the time. A useful sanity check for such cases is discussed in the next section.

1.2 Limiting Cases

In most cases, the validity of a solution for a quantity, whose value is known or can be computed trivially in certain cases, can be verified by substituting those specific cases of parameters into them. Then, the solution should reduce to the same answer if it is indeed correct. In the previous section, we identified the fact that l_{max} should scale with v_0 to reject the hypothesis. Proceeding with a new example, perhaps we forgot whether the acceleration of a block on a smooth, massive inclined plane with an angle of inclination θ is $g \sin\theta$ or $g\cos\theta$. Then, two limiting cases can be considered — namely $\theta \to 0$ and $\theta \to \frac{\pi}{2}$. If $\theta \to 0$, the plane essentially becomes a horizontal ground and the acceleration of the block should be zero. If $\theta \to \frac{\pi}{2}$, the plane is virtually a vertical wall and the acceleration of the block must thus be g. $g\sin\theta$ satisfies both of these criteria while $g\cos\theta$ does not. Thus, the former should be chosen.

Besides choosing limiting cases where a parameter $x \to 0$, $x \to \infty$ or $x \to c$ for some constant c, one can also consider the relative sizes of two parameters x and y. Common limiting cases are $x \gg y$, $x = y$, $x \ll y$, $x > y$ and $x < y$. Here is an example. Consider the simplest Atwood's machine with two masses m_1 and m_2 connected by an inextensible string wrapped around a fixed pulley. Let the accelerations of the masses be a_1 and a_2 respectively, taking the upwards direction to be positive. Suppose we have obtained the following expressions for a_1 and a_2:

$$a_1 = \frac{m_2 - m_1}{m_1 + m_2} g,$$

$$a_2 = \frac{m_1 - m_2}{m_1 + m_2} g.$$

Let us check the validity of our solution. When $m_1 \gg m_2$, a_1 and a_2 should become $-g$ and g respectively as m_1 should experience little influence from m_2 and just undergo free fall while m_2 should rise upwards at g to maintain the length of the string. This is indeed the case. If we reverse the relative magnitudes such that $m_1 \ll m_2$, a_1 and a_2 should become g and $-g$ respectively — a criterion that is also satisfied. When $m_1 = m_2$, we predict that $a_1 = a_2 = 0$ due to the symmetry of the system. This condition is once again fulfilled. The comparisons $m_1 > m_2$ and $m_1 < m_2$ can also be used to check the signs of a_1 and a_2 because the heavier mass should fall downwards as common sense should dictate.

1.3 Physical Principles

1.3.1 *Scaling Arguments*

A famous interview question goes as follows: suppose you were shrunk uniformly into the size of a coin, while retaining your density distribution, and trapped inside a 20cm jar with an open lid; how would you escape the jar?

The answer to this question, which may be surprising at first, is to simply jump out of the jar! An average-sized human can jump 20cm but how can a coin-sized one possibly attain the same height? The crux of this problem is to model how the height of one's jump scales with one's size, which can be quantified by a single length dimension L. It turns out that muscle strength is widely accepted to scale with its cross-sectional area. A crude model behind this property is as follows. The number of muscle fibers is proportional to the cross-sectional area such that a larger cross-sectional area enables the packing of more fibers, which each exert a certain maximum pressure, in parallel. Therefore, the force F that your muscles can exert scales with L^2. Now, to jump a certain height h given that your mass is m, your muscles must do

$$W = mgh$$

amount of work, assuming that the gravitational field strength g is uniform. Now, the work done by your muscles is the force F exerted multiplied by the distance that they contract, d i.e.

$$W = Fd.$$

Therefore,

$$h = \frac{Fd}{mg};$$

d obviously scales with L while m scales with volume and hence, L^3, if the original density distribution (even if it is non-uniform) is retained. Then,

$$h \propto \frac{L^2 \cdot L}{L^3} = \text{constant}.$$

That is, the height that one can jump is independent of the size of the person! The reason behind the enormous sizes of basketball players is not that they can jump higher but because they can easily reach the basket with their long limbs.

The most common scaling arguments are related to the length, area and volume of an object. In general, if all length dimensions of an object are uniformly scaled up by a factor k, its area is scaled up by a factor k^2 while its volume increases by a factor k^3.

Problem: Assuming that the rate of heat loss of a mammal obeys $H = cS$ where c is a constant and S is the total surface area of the mammal, estimate the value of the constant b if $H = aM^b$ where a is a constant and M is the mass of the mammal.

By scaling arguments, if all length dimensions of the mammal are uniformly increased by a factor of k, S increases by a factor of k^2 while M increases by a factor of k^3 if we assume that the density is stretched uniformly. Then, it is evident that $b = \frac{2}{3}$.

In certain cases, we may be given predetermined parameters and thus do not need to model the situation as a quantity of concern can be related to these parameters via dimensional analysis. For example, given that the moment of inertia has dimensions $[M][L]^2$, we can argue that the moment of inertia of a uniform rod of length $2l$ about its centroid is eight times that of a uniform rod of length l and the same mass density, as the moment inertia of a rod should be proportional to its mass multiplied by its squared length (as there is only a single length dimension for a rod). The former rod has twice the mass and twice the length of the latter rod — causing its moment of inertia to be eight times that of the latter.

Problem: An aircraft A travels at a constant altitude while moving at speed u relative to the atmosphere. Determine the speed u' that an aircraft B, with twice the length dimensions of A and all other parameters held constant, must travel at, relative to the atmosphere, for it to remain at a constant altitude. Assume that the lift experienced by an aircraft is independent of the gravitational field strength g.

For an aircraft to remain at a certain vertical height, its weight must be balanced by the lift, F. We can determine how F scales with the length dimension of the aircraft (denoted by l) via dimensional analysis. The possible parameters are the mean density of the atmosphere $\rho_a = [M][L]^{-3}$, the speed of the aircraft relative to the atmosphere $v = [L][t]^{-1}$ and the length dimension of the aircraft $l = [L]$. Suppose

$$F = k\rho_a^\alpha v^\beta l^\gamma$$

for some dimensionless constant k. Since $F = [M][L][t]^{-2}$,

$$\alpha = 1,$$
$$\beta = 2,$$

$$\gamma = 2,$$
$$\implies F = k\rho_a v^2 l^2.$$

We have $F = mg$, where m is the mass of the aircraft, for forces to be balanced in the vertical direction.

$$\implies v = \sqrt{\frac{mg}{k\rho_a l^2}}.$$

As $m \propto l^3$,

$$v \propto \sqrt{l}$$
$$\implies u' = \sqrt{2}u.$$

1.3.2 *Symmetry*

Next, in certain situations which involve some forms of symmetry, some parameters are often "indistinguishable" from each other. Even if those parameters were swapped or permuted, the result should remain unchanged due to symmetry. Then, this fact can be utilized to eliminate options that are not symmetric in those variables.

Consider the simple Atwood's machine with two masses m_1 and m_2 and a fixed pulley again. Let the magnitude of the tension in the string be T. What can we say about the expression for T? Well, T shouldn't depend on whether m_1 is on the left or right. In other words, if the two masses were interchanged such that m_1 becomes m_2 and m_2 becomes m_1, T should not change. Therefore, we can immediately eliminate options such as $T = \frac{m_1}{2}g$ and $T = \frac{m_1 m_2}{m_1 + 2m_2}g$ which are not symmetric in m_1 and m_2. The correct solution for T is

$$T = \frac{2m_1 m_2}{m_1 + m_2}g.$$

We see that even if we interchange m_1 and m_2, T stays the same. Conversely, if the situation is not symmetric, we can also eliminate options that are symmetric. For example, if a particle of mass m_1 travels at a velocity u and undergoes an elastic collision with another initially stationary particle of mass m_2, the expression for the final velocities of the particles should not be symmetric.

Moving on, the symmetry of a system can do much more than eliminate options — it can elucidate the evolution of a system. Systems which start off symmetrical often maintain their symmetry. For example, if a block travels in the positive x-direction at speed v and sticks with an identical block travelling in the negative x-direction at speed v, the velocity of the combined block can only be zero. There is no reason to prefer a single direction in their resultant motion by ascribing them a non-zero final velocity.

1.3.3 *Equivalent Frames*

The combination of symmetry and the notion that the laws of physics hold the same for all inertial frames[1] can lead to enlightening results. Consider the following problem.

Problem: A block travels in the positive x-direction at speed v and sticks with an identical block that was initially stationary. What is the final velocity of the combined block?

We can deduce nothing in the current frame but if we switch to a frame that travels at velocity $\frac{v}{2}$ in the positive x-direction, the situation is symmetrical. Both blocks travel towards each other at speed $\frac{v}{2}$ and must thus have zero final velocity in this frame by the argument in the previous section. Switching back to the original frame, we conclude that the combined block must travel at $\frac{v}{2}$ in the positive x-direction. In fact, we can extend this argument to determine the final velocity due to a perfectly inelastic collision between a block of mass pm and another of mass qm where p and q are integers — you shall do this in Problem 15.

Another common equivalence to exploit is that between accelerating frames and gravity. Simply put, systems cannot distinguish if they are placed in an accelerating frame or a region with a commensurate gravitational field opposite to the direction of acceleration. The field is uniform in space but is possibly mutable in time (as the acceleration of the frame varies).

Problem: Given that a pendulum of length l has angular frequency $\sqrt{\frac{g}{l}}$ in an inertial frame, determine the angular frequency of a pendulum of length l attached to the ceiling of a lift that is accelerating upwards at a constant acceleration a.

The pendulum cannot distinguish if it is in an accelerating lift or a region with gravitational field strength $g+a$. Therefore, its angular frequency must be $\sqrt{\frac{g+a}{l}}$.

1.3.4 *Reversibility*

Finally, most processes in which no heat is generated or transferred are reversible.[2] That is, if a movie of the process is captured and we rewind the tape, the evolution of the process in reverse is perfectly coherent with the physical laws (i.e. it is permitted by the physical laws). Examples of

[1] This will be defined in Chapter 4.

[2] The more formal terminology is actually time reversal symmetry which prevents confusion with the thermodynamic concept of reversibility.

reversible processes are projectile motion and elastic collisions where no heat is generated. One can then exploit the symmetry of such systems in time by considering the reverse process. In the case of projectile motion where a projectile is tossed from and lands on a flat ground, the horizontal and vertical distances covered by the projectile must be identical during the first t seconds and the last t seconds — for the last t seconds of the reversed movie depicts the first t seconds of the original one, with the projectile travelling in the opposite direction horizontally. Furthermore, if we are provided with the additional information that the projectile can only attain a single peak, the horizontal coordinate of the peak must correspond to the midpoint of the starting and ending points of the projectile. The time taken for the projectile to reach the peak from the beginning must also be equal to the time it takes to reach its final location from the peak.

Problem: A particle travelling at velocity u_1 undergoes a one-dimensional elastic collision with an identical particle initially travelling at velocity $u_2 < u_1$ (both velocities are in the same direction). If their final velocities are aligned with their initial velocities and are v_1 and v_2 respectively, show that $v_2 - v_1 = u_1 - u_2$.

Consider a new frame that travels at velocity $\frac{u_1+u_2}{2}$ with respect to the original one. In this frame, particle 1 travels at velocity $\frac{u_1-u_2}{2}$ while particle 2 travels at $-\frac{u_1-u_2}{2}$. By symmetry, particles 1 and 2 must travel at velocities $-v$ and v after the collision. Now, by dimensional analysis, v should be proportional to $\frac{u_1-u_2}{2}$ as that is the only parameter here. Let the dimensionless constant of proportionality be k. Next, by the equivalence of inertial frames, the process in this frame should also be reversible. By considering the reverse process, if particles 1 and 2 had initial velocities v and $-v$, their final velocities should be $-\frac{u_1-u_2}{2}$ and $\frac{u_1-u_2}{2}$ after their collision. Again, $\frac{u_1-u_2}{2}$ must be proportional to v with the same constant of proportionality k for the same physical laws hold in this situation. Since $v = k\frac{u_1-u_2}{2}$ and $\frac{u_1-u_2}{2} = kv$, the only possible solution for k is $k = 1$ (we reject -1 as the particles cannot penetrate each other). Therefore, the final velocities of the particles in this new frame are $-\frac{u_1-u_2}{2}$ and $\frac{u_1-u_2}{2}$. The final relative velocity in the original frame is equal to that in this frame. Thus,

$$v_2 - v_1 = \frac{u_1 - u_2}{2} - \left(-\frac{u_1 - u_2}{2}\right) = u_1 - u_2.$$

Problems

Dimensional Analysis

1. *Cooking Time**

Suppose that the time taken to cook a piece of meat of mass m obeys

$$t = \frac{\alpha c \rho^n m^{\frac{2}{3}}}{k}$$

where α is a dimensionless constant, ρ is the density of the meat, c is the specific heat capacity of the meat and k is the thermal conductivity of the meat. Determine the value of the constant n.

2. *Projectile Motion**

An experiment is performed on both the surfaces of the Earth and Moon where a ball is tossed vertically upwards at an initial velocity v_0. If the time taken by the ball to reach its peak is t_E, maximum height attained by the ball is h_E and the average speed of the ball between its departure from and arrival on the ground is v_E, determine the corresponding values t_M, h_M and v_M for the experiment conducted on the Moon in terms of t_E, h_E and v_E via dimensional analysis.

3. *Throwing Off an Inclined Plane**

A ball is thrown from a long inclined plane with an angle of inclination θ at velocity v_1 which makes an angle ϕ with the slope. When the ball lands on the inclined plane, its final velocity makes an angle α_1 with the slope. Now, repeat the experiment with a new velocity v_2 while maintaining θ and ϕ. Compare the magnitudes of α_1 and α_2.

Limiting Cases

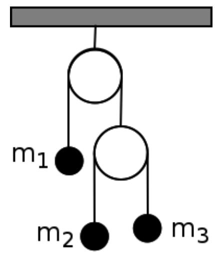
4. *Rowing a Boat**

A boat travels at velocity v_1 relative to still waters. Suppose that a fisherman now decides to cross a river bank with water flowing at speed v_2 downstream. If the boat steers perpendicularly to the velocity of the stream, it reaches the opposite bank after time t_1. If the boat steers at a certain angle upstream such that it reaches the point directly opposite of its starting point, the time taken is t_2. Determine $\frac{v_1}{v_2}$ out of the possible options listed below.

A. $\frac{t_1}{\sqrt{t_2^2 + t_1^2}}$ B. $\frac{t_2}{\sqrt{t_2^2 + t_1^2}}$ C. $\frac{t_1}{\sqrt{t_2^2 - t_1^2}}$ D. $\frac{t_2}{\sqrt{t_2^2 - t_1^2}}$

5. *Stationary Atwood's Machine**

Determine a special combination of m_1, m_2 and m_3 that causes the Atwood's machine depicted below, to be stationary. All strings and pulleys are massless and frictionless.



6. *Perimeter of Circle***

Suppose that you forgot the formula for the perimeter of a circle. By considering the perimeter of a regular N-gon (polygon with N sides) and a particular limiting case, determine the perimeter of a circle with radius r. You will need a special case of L'Hospital's rule which states that if $\lim_{x\to c} f(x) = 0$, $\lim_{x\to c} g(x) = 0$ and $\lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists, $\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$.

Scaling Arguments

7. *Cooling Time**

Estimate how the time taken for a person to cool down after a run scales with the size of the person. Ignore any heat generated by the body after the run and assume that the power radiated from the skin of a person is much larger than the power due to the heat conduction within the person's body. Note that Fourier's law of conduction states that the power delivered by conduction between two surfaces is proportional to their contact area A and temperature gradient.

8. *Maximum Velocity***

Estimate how the maximum sprinting velocity of a mammal on frictionless ice scales with the size of the mammal. Assume that the magnitude of the drag force on the mammal is proportional to its squared speed, $F_{drag} = -bv^2$ where b is independent of v. In determining how b scales with the size of the mammal, it may be helpful to know that the physical cause of the quadratic drag is the bombardment of air molecules on the mammal. What about the maximum velocity that a mammal can stably maintain?

Physical Principles

9. *Bouncing Off a Wall**

A particle is thrown at a vertical wall at a purely horizontal initial velocity. It travels horizontal distance d_1 before undergoing an elastic collision with the wall which causes its horizontal velocity to reverse and its vertical velocity to remain unchanged. The particle then covers horizontal distance d_2 before reaching the ground. Determine the horizontal distance that the particle would have travelled before touching the level ground if the wall were absent.

10. *Projectile Motion on Inclined Plane**

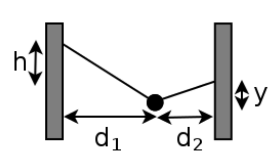
A particle is tossed from an inclined plane of an angle of inclination θ at an initial velocity that is perpendicular to the slope of the plane. Denote the start and end points of the particle's motion, both lying on the plane, to be A and B. In the midst of its motion, the particle attains a maximum perpendicular distance from the plane. Let P denote the foot of the perpendicular from the particle to the plane at this juncture. Determine $\frac{\overline{PB}}{\overline{AP}}$.

11. *Pendulum on Train**

A pendulum of length l is attached to the ceiling of a train which undergoes a constant horizontal acceleration a in the positive x-direction. Given that the pendulum's angular frequency of oscillations in an inertial frame is $\sqrt{\frac{g}{l}}$, determine its equilibrium position and angular frequency on the train.

12. *Pendulum with Strings***

Two strings, which lie in a single plane with the vertical, are attached to a bob as shown in the figure below. The horizontal separations of the end points of the strings with the bob are d_1 and d_2 respectively. The vertical separation between the end points is h and the bob lies a distance y below the right end point. If the bob is given a slight push into the page, determine the angular frequency of small oscillations if the strings remain taut.



13. ***Nearest Points*****

This problem is not really related to physics and is rather contrived but it serves to highlight an important problem-solving technique. Consider a cube of length l with nine particles — eight lie at the vertices of the cube while one lies at the center. Now, each point inside the cube is ascribed to its nearest particle. Determine the total volume of points ascribed to the central particle.

14. ***Right Angles*****

A particle travels at a speed v towards another identical, initially stationary particle, and undergoes a two-dimensional elastic collision where the particles are deflected in different directions. Show that the final velocities of the two particles must be mutually perpendicular.

15. ***Conservation of Momentum*****

It was shown in Section 1.3.3 that a block, initially stationary, approaching an identical, also initially stationary block, at speed v would result in a combined block with final velocity $\frac{v}{2}$ after a perfectly inelastic collision. Now, define m as the mass of a block. When we refer to a block of mass nm where n is an integer, we mean n amalgamated blocks. By using induction, determine the final velocity of the combined block when a block of mass pm, where p is an integer, and initial speed v undergoes a perfectly inelastic collision with an initially stationary block of mass m. Then, determine the final velocity when a block of mass pm and initial velocity v_1 undergoes a perfectly inelastic collision with a block of mass qm, where q is an integer, and initial velocity $v_2 < v_1$ (along the same direction).

Solutions

1. Cooking Time*

Rearranging, we obtain

$$\rho^n = \frac{tk}{\alpha c m^{\frac{2}{3}}}.$$

The dimensions of the thermal conductivity k can be solved for by considering Fourier's law of conduction for one-dimensional heat conduction.

$$\dot{Q} = -kA\frac{dT}{dx},$$

where $A = [L]^2$ represents the contact area, $\frac{dT}{dx} = [T][L]^{-1}$ is the temperature gradient and $\dot{Q} = [E][t]^{-1}$ is the power delivered, where $[E]$ and $[T]$ represent dimensions of energy and temperature respectively. Thus, k has dimensions of $[E][L]^{-1}[t]^{-1}[T]^{-1}$. The dimensions of the specific heat capacity c can be obtained by considering the equation

$$\Delta Q = mc\Delta T,$$

where $\Delta Q = [E]$ is the change in internal energy, $m = [M]$ is the mass of the substance and $\Delta T = [T]$ is the change in temperature.

$$\implies c = \frac{Q}{m\Delta T} = [E][M]^{-1}[T]^{-1}.$$

The expression for ρ^n has dimensions of

$$\rho^n = \frac{[t] \cdot [E][L]^{-1}[t]^{-1}[T]^{-1}}{[E][M]^{-1}[T]^{-1} \cdot [M]^{\frac{2}{3}}} = [M]^{\frac{1}{3}}[L]^{-1}.$$

Hence, $n = \frac{1}{3}$ since $\rho = [M][L]^{-3}$.

2. Projectile Motion*

The possible parameters are the mass of ball $m = [M]$, initial velocity $v_0 = [L][t]^{-1}$ and gravitational field strength $g = [L][t]^{-2}$. By dimensional analysis, the time taken by the ball t to reach its peak should scale as follows:

$$t \propto \frac{v_0}{g}.$$

The surface gravitational field strength of the Moon is one-sixth of the Earth's. Then,

$$t_M = 6t_E.$$

The maximum height attained by the ball h is by dimensional analysis,

$$h \propto \frac{v_0^2}{g}$$
$$\implies h_M = 6h_E.$$

Finally, the average speed of the ball $\langle v \rangle$ can only be proportional to v_0.

$$\langle v \rangle \propto v_0$$
$$\implies v_M = v_E.$$

3. Throwing Off an Inclined Plane*

The only possible parameters are the velocity of the ball $v = [L][t]^{-1}$, the gravitational field strength $g = [L][t]^{-2}$, the mass of the ball $m = [M]$, θ and ϕ. We observe that the expression for α should be independent of v as there is no way to cancel both the length unit $[L]$ and the inverse time unit $[t]^{-1}$ of velocity solely by using g. By a similar argument, α should also be independent of g and m. Therefore, $\alpha_1 = \alpha_2$ as the other parameters, θ and ϕ, are fixed. Finally, this result agrees with the limiting case $\theta = 0$ where $\alpha = \phi$ based on symmetry. Note that one cannot use the limiting case where $\theta = \frac{\pi}{2}$ in this case as the ball will not land on the slope (as it is a vertical wall).

4. Rowing a Boat*

The answer is option D. We can first consider the limiting case where $v_2 = 0$ while v_1 is finite. Then, $\frac{v_1}{v_2}$ tends to infinity. When $v_2 = 0$, it is evident that $t_1 = t_2$. Thus, the only possible options are C and D. To differentiate between these, notice that v_1 must be larger than v_2 for the second scenario to be possible (as it must have a component v_2 upstream). However, option C suggests the possibility that when t_1 is small and $t_2 \gg t_1$ (i.e. $v_1 \approx v_2$ in reality), $\frac{v_1}{v_2} < 1$. Since this is impossible, we eliminate option C and choose D.

5. Stationary Atwood's Machine*

Observe that if $m_2 = m_3 = m$, m_2 and m_3 are indistinguishable, they thus should both remain stationary due to symmetry. The tension in each of the strings holding them is then mg to nullify their weight. For the force on the system comprising the massless pulley and massless string connecting m_2 and m_3 to be balanced, the tension in the string wrapped around the top pulley must be $2mg$. Then, if $m_1 = 2m$ such that its weight balances the tension $2mg$, the entire set-up is stationary.

6. Perimeter of Circle**

Let l denote the length of a line connecting the center of a regular N-gon to one of its vertices. The angle subtended by a line connecting the center to a vertex and one connecting the center to the midpoint of an adjacent edge is $\frac{\pi}{N}$. Therefore, half the length of an edge is $l \sin\frac{\pi}{N}$. Since there are N sides, the perimeter of a regular N-gon is

$$2Nl\sin\frac{\pi}{N}.$$

The perimeter of a circle is then obtained from taking the limit $N \to \infty$ ($l = r$ in this case).

$$\lim_{N\to\infty} 2Nr\sin\frac{\pi}{N} = 2\pi r \lim_{N\to\infty}\frac{N}{\pi}\sin\frac{\pi}{N} = 2\pi r\lim_{x\to 0}\frac{\sin x}{x},$$

where $x = \frac{\pi}{N}$. To compute $\lim\limits_{x\to 0}\frac{\sin x}{x}$, we apply L'Hospital's rule:

$$\lim_{x\to 0}\frac{\sin x}{x} = \lim_{x\to 0}\frac{\cos x}{1} = 1.$$

Therefore, the perimeter of a circle is

$$2\pi r.$$

Another method is to note that for large N, $\frac{\pi}{N}$ is small such that $\sin\frac{\pi}{N} \approx \frac{\pi}{N}$ by the Maclaurin series of $\sin x$. Thus, $\lim\limits_{N\to\infty} 2Nr\sin\frac{\pi}{N} = \lim\limits_{N\to\infty} 2Nr\cdot\frac{\pi}{N} = 2\pi r$.

7. Cooling Time*

Our objective is to determine how the time taken for a person's body to cool from a fixed starting temperature (after running) to his or her normal body temperature scales with size. The limiting factor in this case is the rate of heat conduction within the body as the heat carried to the skin is dispersed at a much faster rate. Let L be a measure of a person's length dimension. Then the area of contact A scales with L^2. Though the person does not have a uniform temperature, let T be a measure of the instantaneous temperature of the person as we will only be looking at how the temperature scales. If Q is the instantaneous internal energy of the person, $Q \propto TL^3$ assuming that the specific heat capacity of the body is constant as the volume of the person scales with L^3. To clarify the dependence of Q on T, the above statement means that if the temperature is doubled everywhere within the body, the stored internal energy will double. Moving on, the rate of heat conduction $\dot{Q}$ is proportional to the product of the contact area A and the temperature gradient — the latter is proportional to $\frac{T}{L}$. Thus,

$$\dot{Q} \propto -L^2 \cdot \frac{T}{L} \propto -L^2 \cdot \frac{Q}{L^4} = -\frac{Q}{L^2}$$
$$\implies \frac{Q}{-\dot{Q}} \propto L^2$$

where a negative sign indicates that a hotter body loses heat. Since $\dot{Q} \propto \dot{T}L^3$,

$$\frac{T}{-\dot{T}} \propto L^2.$$

As different people are presumed to start and end with identical temperatures, the time taken for a person to cool down should scale with his or her squared length dimension as the above expression implies that the time taken for a person's temperature to be reduced by a unit amount scales with L^2, for any given temperature.

8. Maximum Velocity**

When the sprinting velocity is maximum, the power P delivered by the ground to the mammal is equal to the power of the drag force.

$$P = bv^3.$$

Let L be a measure of the mammal's length dimension. b scales with the cross-sectional area of the mammal as the volume rate of air molecules swept by the mammal is proportional to the cross-sectional area (we do not consider v as the dependence of the drag force on v has already been extracted into the v^2 relationship). Next, we move on to the left-hand side. The maximum power delivered by the ground is the maximum force F that the mammal can exert on the ground (by Newton's third law) multiplied by its speed.

$$Fv = bv^3$$
$$v = \sqrt{\frac{F}{b}}.$$

F scales with L^2 for most mammals. Therefore,

$$v \propto \sqrt{\frac{L^2}{L^2}} = \text{constant}.$$

The maximum sprinting velocity should be approximately uniform across mammals. In modeling the stable running speed of a mammal, there are two approaches that we can take — biological and thermodynamic. Biologically, the stable power that a mammal can deliver should be proportional to the volume rate of air intake by the mammal (we assume that oxygen can be instantaneously transported throughout the mammal). This is equal to the cross-sectional area of the windpipe multiplied by the speed of air through the windpipe v_{air}. v_{air} can be estimated by assuming that the lung becomes a perfect vacuum when it expands. Then, by Bernoulli's principle,

$$p_{atm} = \frac{1}{2}\rho v_{air}^2,$$

where p_{atm} is the constant atmospheric pressure and ρ is the density of air in the windpipe which should also be approximately uniform across mammals. Then, v_{air} is constant. The volume rate of air intake is thus proportional to L^2 as the cross-sectional area of the windpipe should scale with L^2. All-in-all, for aerobic sustainability,

$$P \propto L^2$$

$$v = \sqrt[3]{\frac{P}{b}} \propto \sqrt[3]{\frac{L^2}{L^2}} = \text{constant}.$$

In addition to aerobic sustainability, the mammal must also be able to expel heat at a sufficient rate. We can assume that the heat generated by the mammal is proportional to the mechanical power delivered by the mammal (i.e. constant efficiency). Fourier's law of conduction states that the power delivered between two surfaces is proportional to their area of contact and the temperature gradient between them. The temperatures of mammals should not differ by much so the latter scales with $\frac{1}{L}$ while the former scales with L^2. Overall, the power delivered towards the skin of the mammal (where it is radiated) is proportional to L. The stable mechanical power that a mammal can deliver should hence be proportional to L in light of both aerobic and thermodynamic sustainability (the latter is the limiting factor).

$$P \propto L$$

$$v_{stable} = \sqrt[3]{\frac{P}{b}} \propto \frac{1}{\sqrt[3]{L}}.$$

9. Bouncing Off a Wall*

Since the particle simply reverses its horizontal velocity after colliding with the wall, the motion of the particle thereafter is equivalent to that of a particle, located at a horizontal distance d_1 behind the wall and the same vertical height as the initial location of the actual particle (i.e. reflection about the wall), thrown at the same initial speed towards the wall. Therefore, the total horizontal distance covered by this virtual particle (assuming that it penetrates the wall) is $d_1 + d_2$. This must also be the distance covered by the original particle if it is not blocked by the wall due to symmetry (the virtual particle is simply the original particle with a reversed initial velocity).

10. Projectile Motion on Inclined Plane*

We choose our x and y-coordinate axes to be parallel and perpendicular to the slope respectively. Choosing the positive x-direction to be down the slope, the particle experiences an acceleration $(g \sin\theta, g\cos\theta)$. Now, consider a new frame that accelerates at $(g\sin\theta, 0)$ with respect to the original one. In this frame, the particle experiences an acceleration $(0, g\cos\theta)$ — it effectively lives in a world with a modified gravitational field strength $g\cos\theta$ and undergoes a one-dimensional projectile motion (akin to tossing a ball directly upwards). The times taken by the particle to reach its peak (the point where it attains the greatest perpendicular distance from the slope) from the start and to drop back to the slope are identical due to the reversible nature of projectile motion. Let this common time be t. Then, the particle is in the air for $2t$. Now, we analyze the motion of the new frame with respect to the original one to determine the motion of the particle along the slope. In the first t time interval, the particle would have covered $\overline{AP} = \frac{1}{2}g\sin\theta t^2$ horizontal distance as its horizontal acceleration is $g\sin\theta$. In the time interval between t and $2t$, it covers horizontal distance $\overline{PB} = \frac{1}{2}g\sin\theta(2t)^2 - \frac{1}{2}g\sin\theta t^2 = \frac{3}{2}g\sin\theta t^2$ (one can argue that the distance travelled along the slope from the start of the motion should be proportional to t^2 by dimensional analysis too). Thus,

$$\frac{\overline{PB}}{\overline{AP}} = 3.$$

11. Pendulum on Train*

The pendulum effectively experiences an additional uniform and constant horizontal gravitational field $-a$. In order for the pendulum to be at equilibrium, the string must be aligned with the net gravitational field as the tension exerted by a string must be directed along the string. Therefore, the equilibrium position is at a clockwise angle $\tan^{-1}\frac{a}{g}$ from the vertical. Since the net effective gravitational field is $\sqrt{g^2+a^2}$, the angular frequency of the pendulum is $\sqrt{\frac{\sqrt{g^2+a^2}}{l}}$.

12. Pendulum with Strings**

If the strings are to remain taut, the bob has to maintain the same distance with respect to the end points at all times (i.e. the strings are rigid). Therefore, the bob can only rotate about the straight axis that joins the two end points. Let θ be the angle subtended by this axis and the horizontal. The component of gravity currently perpendicular to this axis can be shown to be $g\cos\theta$ by simple trigonometry and is the effective gravity. The component parallel to the axis does not matter as it gets absorbed by the tensions in the strings. Finally, the perpendicular distance between the bob and the axis, which is the effective length of the pendulum, can be shown to be $(y+\frac{hd_2}{d_1+d_2})\cos\theta$. Then, the angular frequency of small oscillations is

$$\omega = \sqrt{\frac{g\cos\theta}{\left(y+\frac{hd_2}{d_1+d_2}\right)\cos\theta}} = \sqrt{\frac{g(d_1+d_2)}{y(d_1+d_2)+hd_2}}.$$

13. Nearest Points**

The trick here is to repeatedly duplicate the original cube into an infinite system of identical cubes with faces stuck together. Each particle in this system has four nearest neighbours that are $\frac{\sqrt{3}l}{2}$ away. Then, we cannot distinguish between the particles at the vertices of the cube and the central particle. That is, the volumes of points ascribed to them should be identical. Let this common volume be V. Notice that each of the particles located at the vertices of the original cube are now sandwiched between eight cubes — the volume inside the original cube ascribed to each vertex is then $\frac{V}{8}$. Equating the ascribed volume V of the central particle and $\frac{V}{8}$ of each of the vertices with the total volume of the original cube,

$$V + \frac{V}{8} \cdot 8 = l^3$$
$$V = \frac{l^3}{2}.$$

14. Right Angles**

Define the velocity of the first particle to be along the positive x-axis. Consider a new frame that travels at velocity $\frac{v}{2}$ with respect to the original frame in the positive x-direction. In this frame, particle 1 travels at velocity $\frac{v}{2}$ while particle 2 travels at velocity $-\frac{v}{2}$. By symmetry, after the collision in this frame, the velocities of particles 1 and 2 must be $-\boldsymbol{u}$ and $\boldsymbol{u}$ for some vector $\boldsymbol{u}$. In order for the process to be reversible in this frame, $u = \frac{v}{2}$ for the same reason as argued in Section 1.3.4. The final velocities in the original frame are

$$\boldsymbol{v}_1 = \frac{\boldsymbol{v}}{2} - \boldsymbol{u}$$
$$\boldsymbol{v}_2 = \frac{\boldsymbol{v}}{2} + \boldsymbol{u}.$$

Taking the dot product of these velocities,

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = \frac{1}{4}\boldsymbol{v} \cdot \boldsymbol{v} - \frac{1}{2}\boldsymbol{v} \cdot \boldsymbol{u} + \frac{1}{2}\boldsymbol{v} \cdot \boldsymbol{u} - \boldsymbol{u} \cdot \boldsymbol{u} = \frac{1}{4}v^2 - u^2 = 0,$$

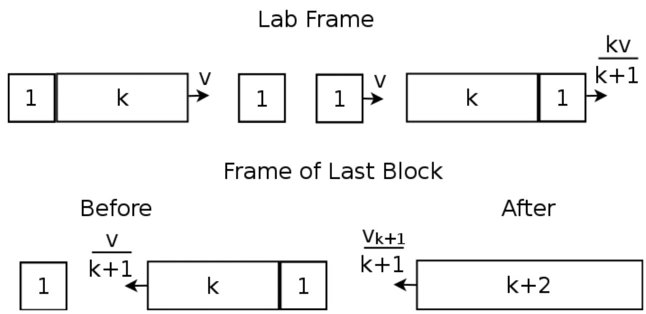
since $u = \frac{v}{2}$. This shows that the two final velocities must be perpendicular.

15. Conservation of Momentum**

Let proposition P_n be that the final velocity v_n of the combined block when a block of mass nm and initial speed v undergoes a perfectly inelastic collision with an initially stationary block of mass m obeys

$$v_n = \frac{n}{n+1}v.$$

All velocities are taken to be positive rightwards. The base case is $n = 1$ which was proven in Section 1.3.3. Given that P_k is true, we shall prove that P_{k+1} is true.



Situation in lab and last block's frames

v_{k+1} is the final velocity of the combined block of mass $(k+2)m$ after a perfectly inelastic collision between a block of mass $(k+1)m$ and another of mass m. The collision process can be split into two parts. Firstly, the k front blocks collide with the lone mass m. By the induction hypothesis, the final velocity of these $(k+1)$ combined blocks is $\frac{kv}{k+1}$. The last mass m at the back of the original $(k+1)$ blocks, which still travels at velocity v, then collides with these $(k+1)$ combined blocks. To determine an alternate expression for v_{k+1} from this procedure, consider the initial frame of the last block (which travels at v relative to the original frame). The combined block of mass $(k+1)m$ now approaches the stationary block of mass m at speed $\frac{v}{k+1}$ leftwards (bottom strip of the figure above). This set-up is identical to the original collision, except with the initial speed scaled down by a factor of $\frac{1}{k+1}$. Therefore, the final velocity of the combined $(k+2)m$ block is $\frac{v_{k+1}}{k+1}$ leftwards by scaling arguments. Switching back to the original frame, the combined $(k+2)m$ blocks travel at velocity $v - \frac{v_{k+1}}{k+1}$ rightwards. Equating this alternate expression for v_{k+1} with v_{k+1},

$$v - \frac{v_{k+1}}{k+1} = v_{k+1}$$
$$v_{k+1} = \frac{k+1}{k+2}v = \frac{k+1}{(k+1)+1}v,$$

which completes our induction. Since P_1 is true, P_n is true for all positive integers n.

$$v_n = \frac{n}{n+1}v.$$

Moving on to the next part, consider the initial frame of the block with mass qm (which travels at v_2 relative to the original frame). The block with mass pm approaches it with velocity $v_1 - v_2$. Now, when pm collides with qm, we can split the collision process into q parts. Firstly, pm collides with the front-most block of qm to form a combined block of mass $(p+1)m$ before it collides with the second block to form a combined block of mass $(p+2)m$ and so on until it collides with the block at the back of the original q blocks. Since $v_n = \frac{n}{n+1}v$ from the previous part, the first collision causes the initial velocity to be scaled by a factor $\frac{p}{p+1}$ while the second causes it to be scaled by a factor $\frac{p+1}{p+2}$ and so on. Therefore, the final velocity of the combined block of mass $(p+q)m$ in the frame that travels at velocity v_2 with respect to the original frame is

$$(v_1 - v_2) \times \frac{p}{p+1} \times \frac{p+1}{p+2} \times \cdots \times \frac{p+q-1}{p+q} = \frac{p(v_1 - v_2)}{p+q}.$$

It is tempting to exploit the fact that scaling the masses of both original blocks by the same factor should not change the final velocity to divide the masses of both blocks by q and apply the previous result (with $n = \frac{p}{q}$) to directly obtain $\frac{p(v_1 - v_2)}{p+q}$ but we cannot do this as the previous result is only valid for integers and not rationals. Having pointed out this flawed argument, we obtain the final velocity in the original frame by adding v_2 to switch back to the original frame.

$$\frac{p(v_1 - v_2)}{p+q} + v_2 = \frac{pv_1 + qv_2}{p+q}.$$

Chapter 2

Infinitesimal Elements

The consideration of infinitesimal elements of a continuous distribution is an essential mathematical procedure in physics as the contribution due to each element to a quantity of interest often differs throughout the whole distribution. For instance, when we wish to calculate the gravitational force on a point mass due to a rod, every mass element on the rod will contribute a force with a different magnitude and possibly, different direction (depending on the orientation). Thus, we must consider infinitesimal elements and use integration to ensure mathematical rigor in such situations.

2.1 One-Dimensional Elements

In many problems, we wish to calculate a quantity, such as the gravitational potential, that is influenced by a distribution of particles with factors that affect it (e.g. mass). Each particle would then contribute a different amount to the quantity being considered. If the distribution is continuous, it is necessary to integrate the contributions over the entire system. Consider the following example.

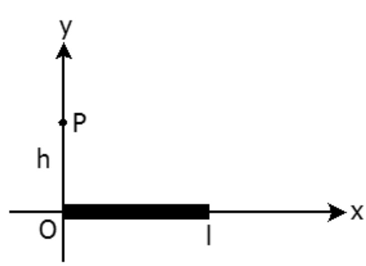


Figure 2.1: Gravitational potential due to a rod

Referring to Fig. 2.1, a uniform thin rod of total mass m and length l lies along the x-axis with ends at $x = 0$ and $x = l$. If the equation for the gravitational potential at a point due to a point particle of mass M is $-\frac{GM}{r}$ where r is the distance of the point of concern from M, determine the gravitational potential due to the rod at point P, $(0, h)$. Now, we consider infinitesimal elements of the rod. Consider a segment with ends at coordinates x and $x + dx$ in Fig 2.2.

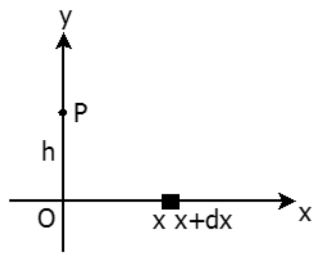


Figure 2.2: Gravitational potential due to a segment between x and $x + dx$

The contribution of this infinitesimal segment of mass $dm = \frac{m}{l}dx$ to the gravitational potential at P is

$$dV = -\frac{Gdm}{\sqrt{h^2 + x^2}} = -\frac{Gm}{l\sqrt{h^2 + x^2}}dx.$$

The total potential due to the rod is obtained by integrating the individual contribution of each infinitesimal element over the entire rod.

$$V = -\int_L \frac{Gm}{l\sqrt{h^2 + x^2}}dx$$

where a symbol L has been written as a subscript of the integral to denote that we are integrating over a one-dimensional line. Furthermore, note that the lower limit of the integral with respect to dV has been taken to be zero as its physical meaning corresponds to the gravitational potential at P due to a rod of zero length. In this case, the rod is a straight line that spans from $x = 0$ to $x = l$. Thus,

$$V = -\int_0^l \frac{Gm}{l\sqrt{h^2 + x^2}}dx.$$

By using the substitution $x = h \tan\theta$, $dx = h \sec^2\theta d\theta$,

$$\begin{aligned} V &= -\int_0^{\tan^{-1}\frac{l}{h}} \frac{Gm}{lh\sec\theta} h\sec^2\theta d\theta \\ &= -\int_0^{\tan^{-1}\frac{l}{h}} \frac{Gm}{l} \sec\theta d\theta \end{aligned}$$

$$= \left[-\frac{Gm}{l} \ln |\sec\theta + \tan\theta| \right]_0^{\tan^{-1}\frac{l}{h}}$$

$$= -\frac{Gm}{l} \ln \left(\sqrt{\frac{l^2}{h^2} + 1} + \frac{l}{h} \right).$$

Now, consider the case where each segment on the rod contributes to a vector quantity at P. For example, what is the gravitational field due to the rod above at P? Note that the gravitational field strength at a point due to a point mass M is $-\frac{GM}{r^2}\hat{\boldsymbol{r}}$ where $\boldsymbol{r}$ is the vector pointing from M to the point of concern.

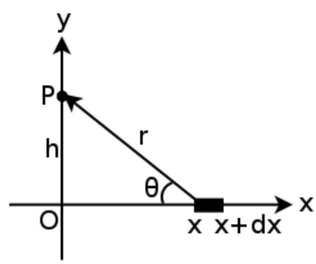


Figure 2.3: Gravitational field due to a segment between x and $x + dx$

Define the xy-plane to be the plane on which the rod and P lie. Now, notice that the unit vector $\hat{\boldsymbol{r}}$ in the case of an infinitesimal element of mass $dm = \frac{m}{l}dx$ between x and $x + dx$ is

$$\hat{\boldsymbol{r}} = \begin{pmatrix} -\cos\theta \\ \sin\theta \end{pmatrix} = \begin{pmatrix} -\frac{x}{\sqrt{h^2+x^2}} \\ \frac{h}{\sqrt{h^2+x^2}} \end{pmatrix},$$

hence the contribution to the gravitational field at P due to this segment is

$$\begin{pmatrix} dg_x \\ dg_y \end{pmatrix} = -\frac{Gmdx}{l(h^2 + x^2)} \begin{pmatrix} -\frac{x}{\sqrt{h^2+x^2}} \\ \frac{h}{\sqrt{h^2+x^2}} \end{pmatrix}.$$

The different components of the total gravitational field at P due to the entire rod can be integrated separately.

$$g_x = \int_0^l \frac{Gmx}{l\,(h^2 + x^2)^{\frac{3}{2}}} dx$$

$$= \left[-\frac{Gm}{l\sqrt{h^2 + x^2}} \right]_0^l$$

$$= \frac{Gm}{lh} - \frac{Gm}{l\sqrt{h^2+l^2}}$$

$$g_y = \int_0^l -\frac{Gmh}{l\left(h^2+x^2\right)^{\frac{3}{2}}} dx$$

$$= \int_0^{\tan^{-1}\frac{h}{l}} -\frac{Gmh}{lh^3\sec^3\theta} h\sec^2\theta d\theta$$

$$= \int_0^{\tan^{-1}\frac{h}{l}} -\frac{Gm}{lh}\cos\theta d\theta$$

$$= \left[-\frac{Gm}{lh}\sin\theta\right]_0^{\tan^{-1}\frac{h}{l}}$$

$$= -\frac{Gm}{l\sqrt{h^2+l^2}}.$$

Now, it just so happened that our distribution in the previous scenario was a straight line. What if the mass distribution took on the form of a more general function $f(x)$ in the xy-plane?

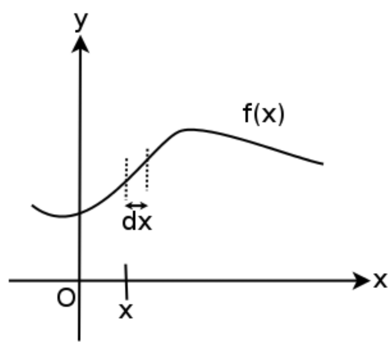


Figure 2.4: Arbitrary curve

Now, we would still need to integrate over individual contributions over the entire curve. However, we cannot simply integrate over x now as this does not reflect the physical curve. Thus, we must consider an appropriate infinitesimal segment to determine what to integrate over. Consider a slanted segment of mass with its ends at x-coordinates x and $x + dx$ (Fig. 2.5). The slanted segment is approximately a straight line in the limit where dx is small.

ds represents the infinitesimal length of the curve while dx and dy merely represent the infinitesimal changes in coordinates. The physical quantity of

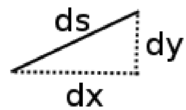


Figure 2.5: General infinitesimal line element

concern is ds as it represents the physical curve.

$$ds = \sqrt{(dx)^2 + (dy)^2} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx. \tag{2.1}$$

Furthermore,

$$\frac{dy}{dx} = f'(x)$$
$$\implies ds = \sqrt{1 + f'(x)^2}\,dx. \tag{2.2}$$

This is the infinitesimal length of an infinitesimal line element in a plane. Suppose that we wish to compute the gravitational potential at the origin due to mass distributed in the form of a curve $f(x)$ with uniform linear mass density λ and ends located at $x = x_0$ and $x = x_1$. Then, the contribution due to an infinitesimal element with its ends at x-coordinates x and $x + dx$ and mass dm is

$$dV = -\frac{Gdm}{\sqrt{x^2 + f(x)^2}}.$$

We can apply the result above to express dm as

$$dm = \lambda ds = \lambda\sqrt{1 + f'(x)^2}\,dx$$

$$V = -\int_{x_0}^{x_1} \frac{G\lambda\sqrt{1 + f'(x)^2}}{\sqrt{x^2 + f(x)^2}}\,dx.$$

Notice that if we had taken $dm = \lambda dx$, we would have obtained an invalid answer as dx does not physically represent an infinitesimal length segment of the curve.

Polar Coordinates

Instead of assigning every point in a two-dimensional space a (x, y) coordinate in a Cartesian system, every point can be instead defined by its distance from the origin and its angular position θ which specifies its orientation in space. Usually θ is defined to be positive anti-clockwise, from an imaginary x-axis. The basis vectors are now the unit vector pointing towards the point

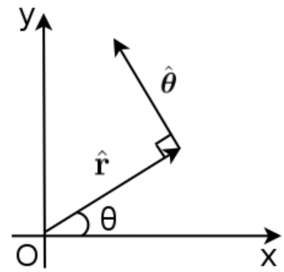


Figure 2.6: Polar coordinates

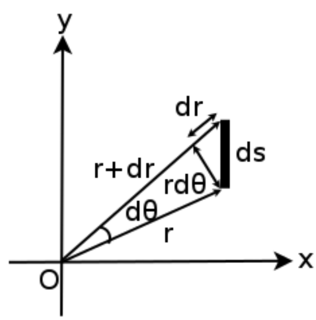


Figure 2.7: Infinitesimal line element in polar coordinates

of concern $\hat{\boldsymbol{r}}$ and a unit vector tangential to it, denoted by $\hat{\boldsymbol{\theta}}$, which points in the direction of increasing θ (Fig. 2.6).

Referring to Fig. 2.7, an infinitesimal length segment now has perpendicular components dr and $rd\theta$ which represent the change in distance from the origin and the infinitesimal arc of radius r that subtends angle $d\theta$.

The infinitesimal length segment is then

$$ds = \sqrt{(dr)^2 + r^2(d\theta)^2} = \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2} d\theta. \tag{2.3}$$

Problem: Compute the total length of the spiral described by the equation $r = \theta$ with limits from $\theta = \theta_0$ to $\theta = \theta_1$.

The total length of a curve is simply the integration of the infinitesimal length segment over the entire curve.

$$S = \int_{\theta_0}^{\theta_1} \sqrt{r^2 + \left(\frac{d\theta}{dr}\right)^2} d\theta = \int_{\theta_0}^{\theta_1} \sqrt{\theta^2 + 1} d\theta.$$

Using the substitutions $\theta = \tan\phi$ and $d\theta = \sec^2\phi d\phi$,

$$S = \int_{\tan^{-1}\theta_0}^{\tan^{-1}\theta_1} \sec^3\phi d\phi.$$

Now, this is a well-known integral which can be solved via the following trick. Using integration-by-parts,

$$\begin{aligned}
I &= \int \sec^3 x dx \\
&= \tan x \sec x - \int \tan^2 x \sec x dx \\
&= \tan x \sec x - \int (\sec^2 x - 1) \sec x dx \\
&= \tan x \sec x + \int \sec x dx - I \\
&= \tan x \sec x + \ln|\tan x + \sec x| - I + c \\
\implies I &= \frac{1}{2}\tan x \sec x + \frac{1}{2}\ln|\tan x + \sec x| + c \\
S &= \frac{1}{2}\left(\theta_1\sqrt{1+\theta_1^2} - \theta_0\sqrt{1+\theta_0^2}\right) + \frac{1}{2}\ln\left|\frac{\theta_1 + \sqrt{1+\theta_1^2}}{\theta_0 + \sqrt{1+\theta_0^2}}\right|.
\end{aligned}$$

2.2 Two-Dimensional Elements

If a distribution spans more than a single dimension, infinitesimal elements of higher dimensions must naturally be considered. In the two-dimensional case, infinitesimal surface elements must be considered. The following are some common ones.

Planar Surface

For a planar surface, a Cartesian system can be adopted. Then, the infinitesimal surface element (Fig. 2.8) is a rectangle of sides dx and dy and area $dA = dxdy$. Cartesian coordinates are especially useful if the region that the distribution spans can be conveniently expressed in terms of x and y. For example, for a rectangle with edges parallel to the x and y-axes, you know that y ranges between two values for every x and the entire range of x ranges between two values.

If the distance of every point on the planar surface to the origin can be expressed neatly in terms of its angular coordinate, θ, it is convenient to

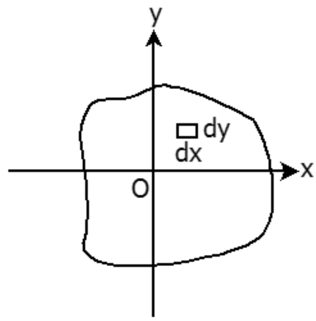


Figure 2.8: Planar surface in Cartesian coordinates

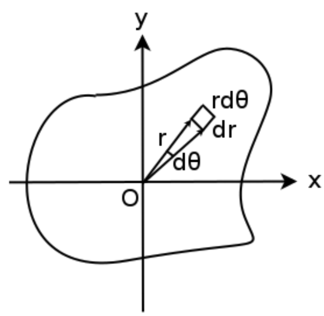


Figure 2.9: Planar surface in polar coordinates

adopt a polar coordinate system. Then, the infinitesimal surface elements are "rectangles" (in the limit of small arc lengths) of sides dr and $rd\theta$ (Fig. 2.9).

Surfaces Formed by Rotating a Curve

A common surface is that formed by rotating a curve $x(z)$ in the region $x \geq 0$ over one complete revolution about the z-axis (Fig. 2.10). Consider a thin trapezoid with circular bases parallel to the xy-plane and at z-coordinates z and $z+dz$, depicted in Fig. 2.11. The radii of the circular bases are $x(z)$ and $x(z+dz) = x+dx$ respectively. The relevant infinitesimal surface between z and $z+dz$ on the original surface is the curved surface of the thin trapezoid.

The most elementary infinitesimal surface element on this surface is a slanted rectangle[1] with area $x(z)d\phi ds$ where ϕ is the azimuthal coordinate in the x-y plane ($x(z)d\phi$ is the length of an infinitesimal arc) and ds is the length of a line along the surface between z and $z+dz$. Now, notice that

[1]Technically, it is a trapezoid with parallel sides of length $x(z)d\phi$ and $x(z+dz)d\phi$ but the difference is third-order after multiplication with ds. Thus, the length of both sides can be taken as $x(z)d\phi$.

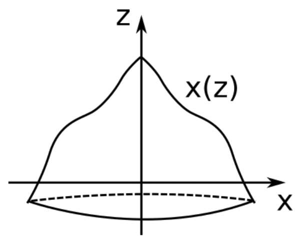


Figure 2.10: Surface formed by rotating curve $x(z)$ about z-axis

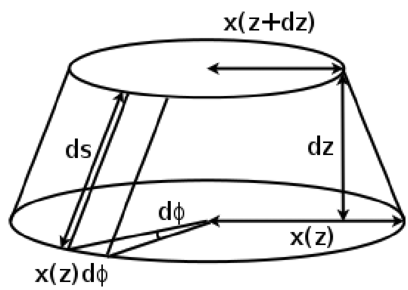


Figure 2.11: Trapezoid between z and $z + dz$

ds is precisely the one-dimensional length segment that has been analyzed before.

$$ds = \sqrt{(dx)^2 + (dz)^2} = \sqrt{\left(\frac{dx}{dz}\right)^2 + 1}\, dz.$$

Therefore, the infinitesimal surface element is

$$dA = x\sqrt{\left(\frac{dx}{dz}\right)^2 + 1}\, dz d\phi.$$

Now that the necessary infinitesimal elements have been studied, we will focus on evaluating the quantities associated with two-dimensional distributions. In general, a double integral will be necessary.

Double Integrals

As its name implies, double integrals imply integrals over two variables. They are naturally required if the distribution spans two dimensions. Double integrals can be evaluated like any normal integral by integrating over one variable layer by layer. However, care must be taken to ensure that the limits of integration indeed represent the entire distribution. Another important fact to understand is that the limits of integration should only depend on the region that we are integrating over and not the integrand. Consider the two following examples.

Problem: Determine the gravitational potential at the origin due to a circular mass distribution centered about the origin with a possibly non-uniform surface mass density $\sigma(r, \theta)$ and radius R. Finally, substitute the special case where $\sigma = \alpha r$ for some constant α.

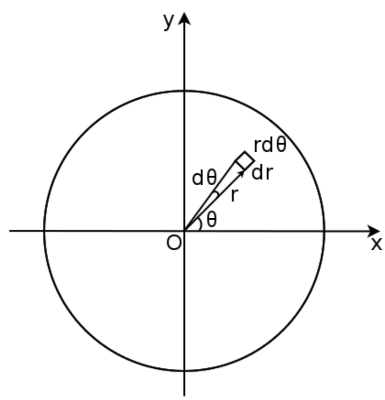


Figure 2.12: Circle in polar coordinates

We have to integrate the contributions due to each infinitesimal surface element. It is natural to adopt polar coordinates in integrating over a circle. Consider an element at coordinates (r, θ). The mass of this element is $\sigma dA = \sigma r d\theta dr$. Hence, its contribution to the potential at the origin is

$$-\frac{G\sigma dA}{r} = -G\sigma d\theta dr.$$

The total gravitational potential is obtained by integrating this quantity over the entire circle:

$$V = -\iint_S G\sigma(r, \theta) d\theta dr.$$

In this case, the limits for r and θ are independent. r ranges from 0 to R while θ ranges from 0 to 2π so

$$V = -\int_0^R \int_0^{2\pi} G\sigma(r, \theta) d\theta dr = -\int_0^{2\pi} \int_0^R G\sigma(r, \theta) dr d\theta.$$

Note that the inner integral is always evaluated first (analogous to the order of brackets). In the first expression, the integral over θ is evaluated before r and vice versa for the second expression. The limits of the integrals do not depend on the order of integration in this case but this is not always true in general. In this particular case, integrating over θ before r would correspond to first computing the contributions due to a thin circular shell of radius r before summing the contributions due to all circular shells from radius 0 to radius R. On the other hand, the reverse order of integration would imply that we first determine the contribution due to a radial line of length R, that subtends a certain angle θ with the x-axis, before summing the contributions due to all lines from $\theta = 0$ to $\theta = 2\pi$. As both orders of integration reflect a circle of radius R, centered about the origin, the exact order of integration does not matter. Substituting the special case where $\sigma = \alpha r$,

$$V = -\int_0^R \int_0^{2\pi} G\alpha r d\theta dr = -\int_0^R 2\pi G\alpha r dr = -\pi G\alpha R^2.$$

Now, consider the more cumbersome distribution below where the limits depend on the order of integration.

Problem: Mass is distributed, with a uniform surface mass density σ, over the area bounded by the x and y axes, and lines $x = l$ and $y = f(x)$. If an infinitesimal surface element at coordinates (x, y) makes a contribution to the quantity of concern given by the expression $\frac{dmx}{M}$ where dm is the mass of the infinitesimal element, and M is the total mass of the distribution that can be taken as a given, compute this quantity of concern — it physically represents the x-coordinate of the center of mass of the distribution.

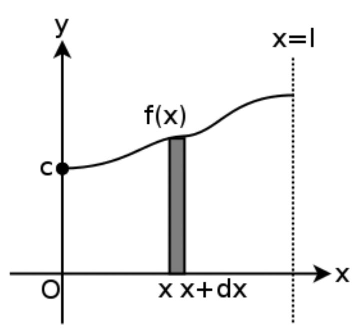


Figure 2.13: Mass under an arbitrary curve

We must integrate individual contributions over the entire surface. To this end, we will consider squares of sides dx and dy in Cartesian coordinates. Integrating the contributions of all such elements over the relevant region,

$$Q = \iint_S \frac{dmx}{M} = \iint_S \frac{\sigma x}{M} dxdy,$$

where S in the subscript denotes the surface of integration. Now, we must take care in writing the limits based on the order of integration. We can first integrate over y and then x. Physically, this represents first finding the contribution due to a vertical strip at a particular x-coordinate (shaded in Fig. 2.13 above) and then integrating over all strips. Then, the limits of the integral over y begin from 0 and end at $f(x)$ for a given x-coordinate. Hence, the limits of integral over y are now a function of x. After integrating over a vertical strip at a particular x-coordinate, we perform an integration over x from 0 to l to sum up the contributions of all strips.

$$Q = \int_0^l \int_0^{f(x)} \frac{\sigma x}{M} dy dx = \int_0^l \frac{\sigma f(x) x}{M} dx.$$

We could have also integrated with respect to x first and then y. This is equivalent to integrating over a horizontal strip first. However, the limits for the x integral would now depend on y and cannot be expressed by a convenient given function. If we absolutely insist on applying this method, the limits for x for a given y in the region $0 \leq y \leq c$ range from $x = 0$ to $x = l$. Moreover, the limits for x for a given y in the region $c \leq y \leq f(l)$ range from $x = f^{-1}(y)$ to $x = l$ where $f^{-1}(y)$ is the inverse function. That is, we now have to split the double integral into two parts and determine the inverse function — a rather troublesome approach.

Let us end this section by analyzing an example of a surface with circular symmetry.

Problem: In Fig. 2.14, consider a spherical cap that is formed by slicing a sphere of radius R by a plane. The base of a spherical cap is a circle with radius a while the altitude from the vertex to the base is h. Determine the area of the curved surface of this spherical cap.

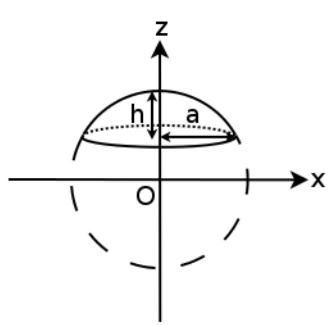


Figure 2.14: Spherical cap

Recall that the infinitesimal element to consider in such cases is a rectangle of sides $ds = \sqrt{(\frac{dx}{dz})^2 + 1}dz$ and $xd\phi$. Since all relevant functions are independent of ϕ, the surface area of a thin trapezoid between z-coordinates z and $z + dz$ in the previous section is

$$2\pi x ds = 2\pi x\sqrt{\left(\frac{dx}{dz}\right)^2 + 1}dz,$$

as the integral over ϕ from 0 to 2π trivially evaluates to 2π. Now we let the center of the original sphere be located at the origin and let the z-axis pass through the vertex of the spherical cap. The radius of the cross-section of the spherical cap $x(z)$ is obtained from the fact that

$$\begin{aligned} x^2 + z^2 &= R^2 \\ x &= \sqrt{R^2 - z^2} \\ \frac{dx}{dz} &= \frac{-z}{\sqrt{R^2 - z^2}}. \end{aligned}$$

The surface area of the spherical cap, A, is obtained via integrating over all trapezoids from $z = R - h$ to $z = R$:

$$\begin{aligned} A &= \int_{R-h}^{R} 2\pi\sqrt{R^2 - z^2} \cdot \sqrt{\frac{z^2}{R^2 - z^2} + 1}dz \\ &= \int_{R-h}^{R} 2\pi R dz \\ &= 2\pi Rh. \end{aligned}$$

Since the question did not give us R, we should express R in terms of a and h. This can be done through the intersecting chords theorem which states that

$$\begin{aligned} a^2 &= h(2R - h) \\ \implies 2Rh &= h^2 + a^2. \end{aligned}$$

Therefore, we have a beautiful result:

$$A = \pi(h^2 + a^2).$$

The surface area of a spherical cap is simply that of a circle with radius $\sqrt{h^2 + a^2}$ (i.e. the length of a line connecting the vertex to a point on the circumference of the base).

2.3 Three-Dimensional Elements

In the three-dimensional case, infinitesimal volume elements should be isolated.

Cartesian Coordinates

In Cartesian coordinates, an infinitesimal box element has lengths dx, dy and dz (Fig. 2.15).

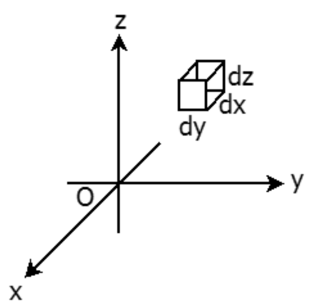


Figure 2.15: Cartesian coordinates

Spherical Coordinates

For integrations over the volume of a sphere, the following spherical coordinate system should be adopted.

In spherical coordinates, every point in space can be defined by its distance to the origin r, the angle θ subtended by its position vector and an imaginary z-axis and ϕ, the angle subtended by the x-axis and the projection of its position vector on to the x-y plane (Fig. 2.16).

The infinitesimal volume element in this case is a small box with sides dr, $rd\theta$ and $r\sin\theta d\phi$. $rd\theta$ is the infinitesimal arc length in the plane containing the z-axis and the position vector $\boldsymbol{r}$. As $r\sin\theta$ is the radius of the circle corresponding to angle θ, labelled as C in Fig. 2.16, $r\sin\theta d\phi$ is the infinitesimal arc length along this circle. Let us familiarize ourselves with the limits of integration over a sphere by calculating the volume of a sphere. Integrating the infinitesimal volume element $dV = rd\theta \cdot r\sin\theta d\phi \cdot dr = r^2 \sin\theta dr d\theta d\phi$,

$$V = \iiint_V r^2 \sin\theta dr d\theta d\phi.$$

Similar to double integrals, a triple integral can be evaluated by integrating from the inner layers while adopting the appropriate limits. Consider the

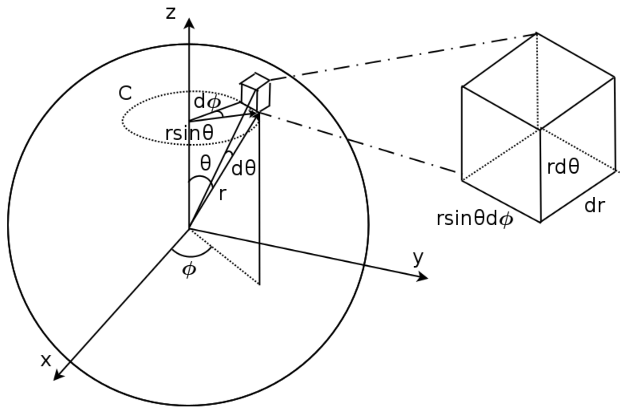


Figure 2.16: Spherical coordinates

scenario where we integrate in the particular order of ϕ, θ and r. In the case of a complete sphere, ϕ ranges from 0 to 2π which is tantamount to integrating over the circumference of C at a fixed r (which is a variable and not the radius of the sphere R) and θ. For fixed r, θ ranges from 0 to π where 0 corresponds to a circle similar to C (but with zero radius) at the pole of a sphere, of radius r, on the positive z-axis while π corresponds to that at the pole on the negative z-axis. Performing the above two integrations would be equivalent to evaluating the contributions due to a thin spherical shell of thickness dr at radius r. To integrate over the volume of the whole sphere, we simply have to integrate r from 0 to R to sum the contributions of shells with radii varying from $r = 0$ to $r = R$. Hence, the volume integral is

$$\begin{aligned}V &= \int_0^R \int_0^\pi \int_0^{2\pi} r^2 \sin\theta d\phi d\theta dr \\ &= \int_0^R \int_0^\pi 2\pi r^2 \sin\theta d\theta dr \\ &= \int_0^R 4\pi r^2 dr \\ &= \frac{4}{3}\pi R^3.\end{aligned}$$

Actually, all permutations of the order of integration in spherical coordinates are valid when integrating over a sphere — a fact that one can check for. Lastly, it is evident that if we are interested in integrating over the surface of the sphere, we can consider infinitesimal rectangular elements of sides $rd\theta$ and $r\sin\theta d\phi$.

Cylindrical Coordinates

Every point on a cylinder can be described by polar coordinates r and θ in the plane of its cross-section and a translational coordinate z along the cylindrical axis.

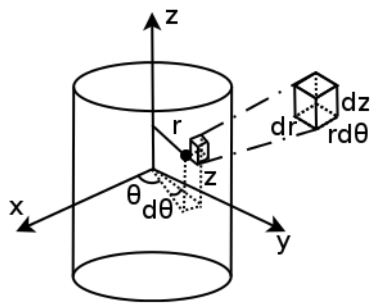


Figure 2.17: Cylindrical coordinates

The infinitesimal volume element in cylindrical coordinates is a "box" of sides dr, $rd\theta$ (arc of radius r that subtends $d\theta$) and dz. The order of integration over r, θ and z again does not matter when it comes to a cylindrical distribution. It can also be seen that if we wish to integrate over the curved surface of a cylinder, we should study infinitesimal elements of sides $rd\theta$ and dz.

Problems

There aren't many problems in this chapter as one would naturally get more practice as one progresses through this book. Rather, the problems in this chapter are intended to exemplify certain common tricks to exploit as well as pitfalls to avoid in integrating.

1. *Line Element in Polar Coordinates**

Prove Eq. (2.3) by expressing the position of a point (r, θ) in Cartesian coordinates and subsequently applying Eq. (2.1).

2. *Pyramid***

Prove that the volume of a pyramid is $\frac{1}{3}Bh$ where B is the area of the base and h is the height of the pyramid.

3. *An Elegant Integral***

Here's a trick in evaluating the integral

$$I_t = \int_{-\infty}^{\infty} e^{-t^2} dt.$$

By considering $I_x \cdot I_y$, where the subscripts x and y indicate the corresponding substitutions for t, in polar coordinates, evaluate I_t.

4. *Another Elegant Integral***

Compute

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 e^{-(x^2+y^2)^2} dx dy.$$

5. *Changing Tricks***

Evaluate

$$\int_1^8 \int_1^{\sqrt[3]{x}} \frac{x}{448y - y^7 - 26} dy dx.$$

6. *Sphere in Cartesian Coordinates***

Write down the limits of integration over a sphere of radius R, centered at the origin, in Cartesian coordinates. Verify that your limits give the correct volume of a sphere.

7. *Potential at Center of Cylinder***

Determine the gravitational potential at the center of a uniform cylinder of radius R, length L and uniform mass density ρ.

8. *Potential due to a Uniform Sphere****

Determine the gravitational potential at a point P that is located a distance s away from the center of a uniform sphere of mass density ρ and radius R. Consider both cases where $s > R$ and $s \leq R$.

Solutions

1. Line Element in Polar Coordinates*

The position of a point (r, θ) in Cartesian coordinates is

$$\begin{aligned} x &= r\cos\theta \\ y &= r\sin\theta \\ dx &= dr\cos\theta - r\sin\theta d\theta \\ dy &= dr\sin\theta + r\cos\theta d\theta \\ ds &= \sqrt{(dx)^2 + (dy)^2} \\ &= \sqrt{(dr\cos\theta - r\sin\theta d\theta)^2 + (dr\sin\theta + r\cos\theta d\theta)^2} \\ &= \sqrt{(dr)^2 + r^2(d\theta)^2} \\ &= \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2} d\theta. \end{aligned}$$

2. Pyramid**

Define the x-axis to pass through the vertex of the pyramid (which is set as the origin) and be parallel to the height of the pyramid. Consider a disk of infinitesimal thickness dx, whose parallel surfaces are perpendicular to the x-axis and are at x-coordinates x and $x + dx$. The two surfaces of the disk have areas A and $A + dA$ respectively and the total volume of this element is

$$\frac{1}{2}(A + A + dA)dx = Adx.$$

Due to the similarity between the entire pyramid and the pyramid from the top up to a distance x from the vertex,

$$\frac{A}{B} = \frac{x^2}{h^2}.$$

The volume is obtained by integrating x from 0 to the height of the pyramid h:

$$V = \int_0^h Adx = \int_0^h \frac{x^2}{h^2} Bdx = \frac{1}{3}Bh.$$

3. An Elegant Integral**

$$\begin{aligned}
I_x \cdot I_y &= \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dxdy \\
&= \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r d\theta dr \\
&= \int_0^{\infty} 2\pi e^{-r^2} r dr \\
&= \left[-\pi e^{-r^2}\right]_0^{\infty} \\
&= \pi
\end{aligned}$$

where we have expressed the integral in terms of polar coordinates in the third equality. Since $I_x = I_y = I_t$,

$$I_t = \sqrt{\pi}.$$

4. Another Elegant Integral**

Let

$$I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 e^{-(x^2+y^2)^2} dxdy.$$

Consider another integral

$$I' = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 e^{-(x^2+y^2)^2} dxdy.$$

By symmetry,

$$I = I'.$$

Adding the two integrals above,

$$2I = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x^2 + y^2) e^{-(x^2+y^2)^2} dxdy$$

$$= \int_0^\infty \int_0^{2\pi} r^2 e^{-r^4} r d\theta dr$$

$$= 2\pi \int_0^\infty r^3 e^{-r^4} dr$$

$$= \frac{\pi}{2} \int_0^\infty e^{-u} du$$

$$= \frac{\pi}{2},$$

where we have adopted polar coordinates en route and the substitution $u = r^4$. Finally,

$$I = \frac{\pi}{4}.$$

5. Changing Tricks**

As hinted by the problem's title, the crux of this integral is to swap the order of integration. Let us first visualize the region that we are integrating over (Fig. 2.18). Plot $y = \sqrt[3]{x}$ on a graph. The region of concern (shaded below) is that bounded by lines $y(x)$, $y = 1$ and $x = 8$.

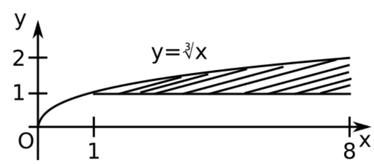


Figure 2.18: Region of integration

If we wish to integrate over x first, we have to integrate the integrand from $x = y^3$ to $x = 8$ for a given y-coordinate (this corresponds to a single horizontal strip between y and $y + dy$). Afterwards, we have to integrate over y from 1 to 2 to sum over all such horizontal strips. Therefore,

$$\int_1^8 \int_1^{\sqrt[3]{x}} \frac{x}{448y - y^7 - 26} dy dx = \int_1^2 \int_{y^3}^8 \frac{x}{448y - y^7 - 26} dx dy$$

$$= \int_1^2 \frac{32 - \frac{y^6}{2}}{448y - y^7 - 26} dy.$$

At this point, we let the denominator be $u = 448y - y^7 - 26$. Then, $du = (448 - 7y^6)dy$ which is 14 times the numerator. Thus, we need to evaluate

$$\int_{421}^{842} \frac{du}{14u} = \frac{1}{14} \ln 2.$$

6. Sphere in Cartesian Coordinates**

Without any loss of generality, let us integrate over a sphere in Cartesian coordinates in the order of x, y and z. For fixed y and z, the limits of integration of x are from $x = -\sqrt{R^2 - y^2 - z^2}$ to $x = \sqrt{R^2 - y^2 - z^2}$ (this corresponds to a chord parallel to the x-axis). For fixed z, y ranges from $y = -\sqrt{R^2 - z^2}$ to $y = \sqrt{R^2 - z^2}$. Performing the two integrations above would be equivalent to integrating over a cross-section of the sphere, perpendicular to the z-axis, at z-coordinate z. Finally, we have to integrate over z from $-R$ to R to sum the contributions from all cross-sections. Consequently, the volume of a sphere can be obtained from integrating $dV = dxdydz$ (a box in Cartesian coordinates) over the region of the sphere.

$$\begin{aligned}
V &= \int_{-R}^{R} \int_{-\sqrt{R^2 - z^2}}^{\sqrt{R^2 - z^2}} \int_{-\sqrt{R^2 - y^2 - z^2}}^{\sqrt{R^2 - y^2 - z^2}} dxdydz \\
&= \int_{-R}^{R} \int_{-\sqrt{R^2 - z^2}}^{\sqrt{R^2 - z^2}} 2\sqrt{R^2 - y^2 - z^2} dydz \\
&= \int_{-R}^{R} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2\sqrt{R^2 - z^2} \cos\theta \cdot \sqrt{R^2 - z^2} \cos\theta d\theta dz \\
&= \int_{-R}^{R} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (R^z - z^2)(\cos 2\theta + 1) d\theta dz \\
&= \int_{-R}^{R} \pi(R^2 - z^2) dz \\
&= 2\pi R^3 - \frac{2\pi R^3}{3} \\
&= \frac{4\pi R^3}{3},
\end{aligned}$$

where we have adopted the substitutions $y = \sqrt{R^2 - z^2} \sin\theta$ and $dy = \sqrt{R^2 - z^2} \cos\theta d\theta$ en route.

7. Potential at Center of Cylinder**

Define the origin at the center. Adopting cylindrical coordinates, an infinitesimal element $dm = \rho r d\theta dr dz$ at (r, θ, z) contributes $-\frac{Gdm}{\sqrt{r^2+z^2}} = -\frac{G\rho r}{\sqrt{r^2+z^2}} d\theta dr dz$ to the potential at the origin as $\sqrt{r^2+z^2}$ is the distance between this element and the origin. The total potential is

$$\begin{aligned} V &= \int_{-\frac{L}{2}}^{\frac{L}{2}} \int_0^R \int_0^{2\pi} -\frac{G\rho r}{\sqrt{r^2+z^2}} d\theta dr dz \\ &= -2\pi G\rho \int_{-\frac{L}{2}}^{\frac{L}{2}} \int_0^R \frac{r}{\sqrt{r^2+z^2}} dr dz \\ &= -2\pi G\rho \int_{-\frac{L}{2}}^{\frac{L}{2}} \left[\sqrt{r^2+z^2}\right]_0^R dz \\ &= -2\pi G\rho \int_{-\frac{L}{2}}^{\frac{L}{2}} \left(\sqrt{R^2+z^2} - z\right) dz. \end{aligned}$$

At this point, we introduce the substitutions $z = R\tan\theta$ and $dz = R\sec^2\theta d\theta$. Furthermore, observing that the integral involving z yields zero as it is an odd function,

$$\begin{aligned} V &= -2\pi G\rho \int_{-\tan^{-1}\frac{L}{2R}}^{\tan^{-1}\frac{L}{2R}} R\sec\theta \cdot R\sec^2\theta d\theta \\ &= -2\pi G\rho R^2 \left(\frac{L}{2R}\sqrt{\frac{L^2}{4R^2}+1} + \ln\left|\frac{L}{2R} + \sqrt{\frac{L^2}{4R^2}+1}\right| \right), \end{aligned}$$

where we have substituted $\int \sec^3 x dx = \frac{1}{2}\tan x \sec x + \frac{1}{2}\ln|\tan x + \sec x| + c$ (computed previously in this chapter). Finally, note that integrating over z before r is much more tedious (always choose the more convenient order of integration).

8. Potential due to a Uniform Sphere***

Define the origin to be at the center of the sphere. Let point P be located on the positive z-axis. Then, the distance between an infinitesimal volume element at spherical coordinates (r, θ, ϕ) and point P is

$$l = \sqrt{s^2 + r^2 - 2rs\cos\theta}$$

by the cosine rule. Note that it does not matter if $s > R$ or $s \leq R$. The contribution by this element $dm = \rho r^2 \sin\theta d\theta d\phi dr$ to the total potential at P is then

$$-\frac{Gdm}{l} = -\frac{G\rho r^2 \sin\theta d\theta d\phi dr}{\sqrt{s^2 + r^2 - 2rs\cos\theta}}.$$

The total potential at P is obtained by integrating the above over the entire sphere.

$$\begin{aligned}
V &= \int_0^R \int_0^\pi \int_0^{2\pi} -\frac{G\rho r^2 \sin\theta}{\sqrt{s^2 + r^2 - 2rs\cos\theta}} d\phi d\theta dr \\
&= \int_0^R \int_0^\pi -\frac{2\pi G\rho r^2 \sin\theta}{\sqrt{s^2 + r^2 - 2rs\cos\theta}} d\theta dr \\
&= \int_0^R \left[-\frac{2\pi G\rho r}{s}\sqrt{s^2 + r^2 - 2rs\cos\theta}\right]_0^\pi dr \\
&= \int_0^R -\frac{2\pi G\rho r}{s}\left(s + r - \sqrt{s^2 + r^2 - 2rs}\right) dr.
\end{aligned}$$

Note that we deliberately integrate over θ before r to greatly simplify the process. At this point, it matters if $s > r$ or $s \leq r$ as it will affect the result from the square root (note that r is a variable and does not refer to the radius of the sphere). If $s > R$, s will be larger than all r as $0 \leq r \leq R$. Then,

$$\begin{aligned}
V &= \int_0^R -\frac{2\pi G\rho r}{s}\left[s + r - (s - r)\right] dr \\
&= \int_0^R -\frac{4\pi G\rho r^2}{s} dr \\
&= -\frac{4G}{3s}\pi\rho R^3 \\
&= -\frac{GM}{s}
\end{aligned}$$

where M is the total mass of the sphere. We see that for $s > R$, the sphere is essentially equivalent to a point mass M at its center. If $s < R$, we have to split the integral over r into two parts — namely from 0 to s (where $r \leq s$)

and from s to R (where $r \geq s$).

$$V = \int_0^s -\frac{2\pi G\rho r}{s}\left[s + r - (s - r)\right] dr$$

$$+ \int_s^R -\frac{2\pi G\rho r}{s}\left[s + r - (r - s)\right] dr$$

$$= \left[-\frac{4\pi G\rho r^3}{3s}\right]_0^s + \left[-2\pi G\rho r^2\right]_s^R$$

$$= \frac{2\pi G\rho s^2}{3} - 2\pi G\rho R^2.$$

Chapter 3

Kinematics

This chapter deals with kinematics — the study of the motion of objects. We will not be working with forces, torques and masses here as we are not concerned about why objects move but rather, the path they take. So, let us pretend that we have not learnt these and instead, focus on the kinematic quantities used to describe motion.

3.1 Vectors

Vectors are crucial mathematical tools in our analyses as they can greatly simplify various formulations. For the sake of our purposes, a vector[1] can be understood as an arrow in space — a line segment with a length and direction. Diagrammatically, vectors are drawn as arrows with an arrow head and a tail. In equations, we shall denote vectors with bolded alphabets, such as $\boldsymbol{A}$.

The utility of vectors stems from their independence from coordinate systems that describe them. For example, if a corporeal arrow is placed in space, different coordinate systems may quantify it in terms of different equations but they all describe the same physical entity! Following from this, translations which maintain the orientation of a vector do not change it. Therefore, the two vectors below are the same.

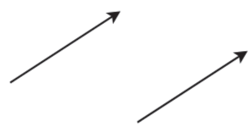


Figure 3.1: Identical vectors

[1] The vectors in this book are generally assumed to have three dimensions, unless stated otherwise.

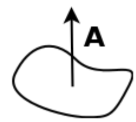


Figure 3.2: Area vector

A concrete example of a vector would be the position vector $\boldsymbol{r}$ of a particle with respect to a certain origin O. The magnitude of $\boldsymbol{r}$, denoted as r, is the length of the straight line connecting O and the particle. $\boldsymbol{r}$ points from O to the particle. By adopting vector notation, the position of a particle can be described without any reference to coordinate axes (but obviously, there must still be an origin that can be used as a reference point).

Another important vector in physics is the area vector of a flat surface. The area vector $\boldsymbol{A}$ is defined to be perpendicular to the surface and its magnitude is equal to the area of the surface (Fig. 3.2).

Observe that there are actually two possible directions for $\boldsymbol{A}$. There is generally no preference[2] for either of them in the case of open surfaces (surfaces which do not enclose any volume), which flat surfaces form a subset of. In the case of closed surfaces — for which an "outside" can be distinguished from an "inside" — the area vectors are defined to be outwards by convention.

Next, a unit vector $\hat{\boldsymbol{a}}$ is a vector whose magnitude is unity and is denoted with a "hat" above its alphabet. For example, the unit vector in the direction of $\boldsymbol{r}$ is

$$\hat{\boldsymbol{r}} = \frac{\boldsymbol{r}}{r},$$

where we have divided by r to ensure that its magnitude is one. Conversely, we can write a vector $\boldsymbol{A}$ in general as

$$\boldsymbol{A} = A\hat{\boldsymbol{A}}. \tag{3.1}$$

3.1.1 *Vector Algebra*

Multiplication by a Scalar

Multiplying a vector $\boldsymbol{A}$ with a scalar c simply scales the original vector by a factor of c.

[2]However, if a line integral is performed over the perimeter of the surface, there is a conventional direction for $\boldsymbol{A}$. This will be elaborated when we reach Ampere's law in magnetism.

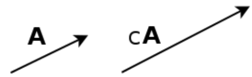


Figure 3.3: Multiplication by a scalar

If $c = -1$, the vector is simply reversed. If c is negative, the vector is first reversed, before it is scaled by a factor of $|c|$.

Addition of Vectors

To add two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$, simply stick the head of one vector to the tail of the other. The dotted lines in Figure 3.4 below depict this procedure. Then, the resultant vector after addition is the vector emanating from the tail of the former and ending at the head of the latter.

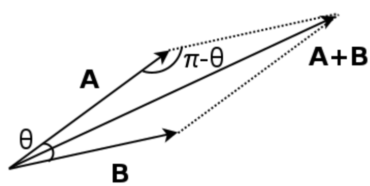


Figure 3.4: Addition of two vectors

If given the angle θ subtended by the two vectors when they are placed head-to-head or tail-to-tail, we can compute the magnitude (i.e. length) of the resultant vector $|\boldsymbol{A}+\boldsymbol{B}|$, where the absolute brackets signify "taking the magnitude of", by applying the cosine rule, i.e.

$$|\boldsymbol{A}+\boldsymbol{B}|^2 = A^2 + B^2 - 2AB\cos(\pi - \theta)$$

$$|\boldsymbol{A}+\boldsymbol{B}| = \sqrt{A^2 + B^2 + 2AB\cos\theta}.$$

To subtract, simply reverse the relevant vector and perform an addition. The magnitude of $|\boldsymbol{A}-\boldsymbol{B}|$ can then be obtained from substituting $\pi + \theta$ (which reverses one vector) for θ in the equation above. An example involving the subtraction of vectors is as follows: suppose that the position vector of a particle is $\boldsymbol{r}$ with respect to origin O. Determine its position vector $\boldsymbol{r'}$ with respect to another origin O' whose position is $\boldsymbol{r_0}$ with respect to O. In this case, $\boldsymbol{r}$, $\boldsymbol{r_0}$ and $\boldsymbol{r'}$ are akin to $\boldsymbol{A}+\boldsymbol{B}$, $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively. Then,

$$\boldsymbol{r} = \boldsymbol{r_0} + \boldsymbol{r'},$$

$$\boldsymbol{r'} = \boldsymbol{r} - \boldsymbol{r_0}.$$

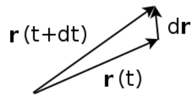


Figure 3.5: Infinitesimal change in a vector

Another example would be the infinitesimal change in a vector (Fig. 3.5). Suppose that $\boldsymbol{r}$ is a function of time t. Then the infinitesimal change in $\boldsymbol{r}$ is

$$d\boldsymbol{r} = \boldsymbol{r}(t+dt) - \boldsymbol{r}(t).$$

Note that the above does not simply represent a change in the magnitude r as the vector could also change in direction. That is,

$$\begin{aligned} d\boldsymbol{r} &= r(t+dt)\hat{\boldsymbol{r}}(t+dt) - r(t)\hat{\boldsymbol{r}}(t) \\ &= [r(t+dt)\hat{\boldsymbol{r}}(t+dt) - r(t+dt)\hat{\boldsymbol{r}}(t)] + [r(t+dt)\hat{\boldsymbol{r}}(t) - r(t)\hat{\boldsymbol{r}}(t)] \\ &= r d\hat{\boldsymbol{r}} + dr\hat{\boldsymbol{r}}. \end{aligned}$$

The first term is associated with the change in direction of $\boldsymbol{r}$ while the second is due to its change in magnitude.

Dot Product

The scalar or dot product of two vectors is an operation that produces a scalar. The dot product $\boldsymbol{A} \cdot \boldsymbol{B}$ is defined as

$$\boldsymbol{A} \cdot \boldsymbol{B} = AB\cos\theta, \tag{3.2}$$

where θ is the angle defined in the section above. Observe that given vectors of fixed magnitudes, their dot product is the largest when they are parallel and zero when they are perpendicular.

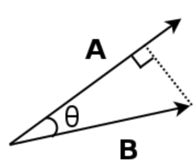


Figure 3.6: Projection of $\boldsymbol{B}$ onto $\boldsymbol{A}$

Notice that $A\cos\theta$ is the length of the projection of $\boldsymbol{A}$ on $\boldsymbol{B}$ and $B\cos\theta$ is the length of the projection of $\boldsymbol{B}$ on $\boldsymbol{A}$. Thus, the geometric interpretation of the dot product of two vectors is the signed product of the magnitude of the projection of one vector on the other, and the magnitude of the other vector. Correspondingly, the dot product of a vector with itself produces its squared magnitude

$$\boldsymbol{A} \cdot \boldsymbol{A} = A^2.$$

Lastly, it can be easily shown that the dot product is both commutative, $\boldsymbol{A} \cdot \boldsymbol{B} = \boldsymbol{B} \cdot \boldsymbol{A}$, and distributive $\boldsymbol{A} \cdot (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \cdot \boldsymbol{B} + \boldsymbol{A} \cdot \boldsymbol{C}$. Due to the distributive property, there is an important identity regarding the derivative of a dot product:

$$\begin{aligned} d(\boldsymbol{A} \cdot \boldsymbol{B}) &= (\boldsymbol{A} + d\boldsymbol{A}) \cdot (\boldsymbol{B} + d\boldsymbol{B}) - \boldsymbol{A} \cdot \boldsymbol{B} \\ &= d\boldsymbol{A} \cdot \boldsymbol{B} + \boldsymbol{A} \cdot d\boldsymbol{B}, \end{aligned}$$

where second order terms have been discarded. The derivative of the dot product of a vector with itself results in an important expression for $\boldsymbol{A} \cdot d\boldsymbol{A}$. Applying the above result,

$$d(\boldsymbol{A} \cdot \boldsymbol{A}) = 2\boldsymbol{A} \cdot d\boldsymbol{A}.$$

Since $d(\boldsymbol{A} \cdot \boldsymbol{A}) = d(A^2) = 2AdA$,

$$\boldsymbol{A} \cdot d\boldsymbol{A} = AdA. \tag{3.3}$$

Problem: Armed with the notion of the dot product, prove the cosine rule for a triangle.

Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{A} + \boldsymbol{B}$ be vectors along the edges of a triangle. Then,

$$\begin{aligned} |\boldsymbol{A} + \boldsymbol{B}|^2 &= (\boldsymbol{A} + \boldsymbol{B}) \cdot (\boldsymbol{A} + \boldsymbol{B}) = A^2 + B^2 + 2\boldsymbol{A} \cdot \boldsymbol{B} \\ &= A^2 + B^2 + 2AB\cos\theta = A^2 + B^2 - 2AB\cos(\pi - \theta), \end{aligned}$$

where $\pi - \theta$ is the angle within the triangle, opposite the edge represented by $\boldsymbol{A} + \boldsymbol{B}$.

Cross Product

The vector or cross product of two vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ results in a vector $\boldsymbol{C}$ which is perpendicular to the plane containing $\boldsymbol{A}$ and $\boldsymbol{B}$. The magnitude of $\boldsymbol{C}$ is defined as

$$C = AB\sin\theta, \tag{3.4}$$

with the same notation of θ as above. The geometric interpretation of $\boldsymbol{C}$ is the area vector of the parallelogram formed with $\boldsymbol{A}$ and $\boldsymbol{B}$ as its sides.

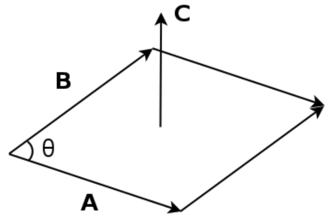


Figure 3.7: Cross product

To select the particular direction of $\boldsymbol{C}$ from the two possible options, we use the following right-hand grip rule. To determine the direction of $\boldsymbol{A} \times \boldsymbol{B}$, first point the four fingers of your right hand in the direction of $\boldsymbol{A}$. Now, curl your fingers towards the direction of $\boldsymbol{B}$. If this can be accomplished after subtending an angle less than π radians, your straightened thumb will point in the direction of the resultant vector $\boldsymbol{C}$.

Due to the assignment of a direction, the cross product is not commutative. In fact, $\boldsymbol{A}\times\boldsymbol{B} = -\boldsymbol{B}\times\boldsymbol{A}$. However, the cross product is still distributive, $\boldsymbol{A} \times (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \times \boldsymbol{B} + \boldsymbol{A} \times \boldsymbol{C}$.

Triple Products

From the definitions of the dot and cross products, it can also be proven that

$$\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B} \cdot (\boldsymbol{C} \times \boldsymbol{A}) = \boldsymbol{C} \cdot (\boldsymbol{A} \times \boldsymbol{B}). \tag{3.5}$$

The geometric meaning of this is the volume of a parallelepiped with those three vectors as its sides. For example, $\boldsymbol{A} \times \boldsymbol{B}$ is the area vector of the base bounded by $\boldsymbol{A}$ and $\boldsymbol{B}$, while taking the dot product of this with $\boldsymbol{C}$ simply multiplies the magnitude of the component of $\boldsymbol{C}$ perpendicular to the base (i.e. the height) with the area of the base.

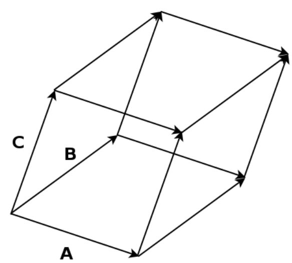


Figure 3.8: Triple product

Another useful identity is

$$\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) = \boldsymbol{B}(\boldsymbol{A} \cdot \boldsymbol{C}) - \boldsymbol{C}(\boldsymbol{A} \cdot \boldsymbol{B}). \tag{3.6}$$

This is commonly known as the "BAC-CAB" rule which is a neat mnemonic. Take note that the cross product is not associative. That is, $\boldsymbol{A} \times (\boldsymbol{B} \times \boldsymbol{C}) \neq (\boldsymbol{A} \times \boldsymbol{B}) \times \boldsymbol{C}$.

Vectors in Coordinate Systems

So far, our analysis has been independent of a coordinate system. To actually extract any meaning from vector operations, a coordinate system is necessary though the concept it describes is actually independent of any coordinate system.

A spatial, three-dimensional coordinate system essentially establishes three basis vectors. Before we explore the definition of a basis vector, we introduce the concept of a linear combination. Suppose that we have n vectors which range from $\boldsymbol{e}_1$ to $\boldsymbol{e}_n$. A general linear combination of this set of vectors is

$$c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2 + \cdots + c_n\boldsymbol{e}_n,$$

where c_1 to c_n are arbitrary scalars. By definition, basis vectors are linearly independent[3] — that is, one cannot be expressed in terms of a linear combination of the others. Furthermore, for a set of three vectors to qualify as a basis for three-dimensional space, every three-dimensional vector must be able to be expressed as a linear combination of the three vectors, that is,

$$\boldsymbol{r} = r_1\boldsymbol{e}_1 + r_2\boldsymbol{e}_2 + r_3\boldsymbol{e}_3,$$

for some scalars r_1, r_2 and r_3, for all $\boldsymbol{r}$ in three-dimensional space. Furthermore, this representation is unique.[4] r_1, r_2 and r_3 are known as the components of a vector. For the sake of convenience, most, if not all, coor-

[3]To check for linear independence, one has to show that $c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2 + \cdots + c_n\boldsymbol{e}_n = \mathbf{0}$– where $\mathbf{0}$ is the n-dimensional null vector–only has the trivial solution $c_1 = c_2 = \cdots = c_n = 0$. If this is violated, there is at least one c_i that is non-zero. Move that term (e.g. $c_k\boldsymbol{e}_k$) to the right-hand side such that $c_1\boldsymbol{e}_n + \cdots + c_n\boldsymbol{e}_n - c_k\boldsymbol{e}_n = c_k\boldsymbol{e}_k$ where the left-hand side excludes $c_k\boldsymbol{e}_k$. Evidently, this shows that $\boldsymbol{e}_k$ can be expressed as a linear combination of the other vectors.

[4]This uniqueness stems from the linear independence property of the basis vectors. Suppose $\boldsymbol{r} = r_1\boldsymbol{e}_1 + r_2\boldsymbol{e}_2 + r_3\boldsymbol{e}_3$ and $\boldsymbol{r} = r_1'\boldsymbol{e}_1 + r_2'\boldsymbol{e}_2 + r_3'\boldsymbol{e}_3$. Then, $(r_1 - r_1')\boldsymbol{e}_1 + (r_2 - r_2')\boldsymbol{e}_2 + (r_3 - r_3')\boldsymbol{e}_3 = \mathbf{0}$ whose only solution is $r_i = r_i'$ for $i = 1, 2, 3$ by the linear independence of the basis vectors (see previous footnote).

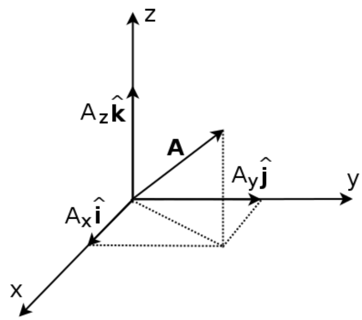


Figure 3.9: Components of a vector in Cartesian coordinates

dinate systems in physics define basis vectors that are mutually orthogonal (that is, $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0$ if $i \neq j$) and are unit vectors. Then, the ith component of a vector can be extracted by taking the dot product of the corresponding basis vector with $\boldsymbol{r}$:

$$\boldsymbol{r} \cdot \boldsymbol{e}_i = r_1 \boldsymbol{e}_1 \cdot \boldsymbol{e}_i + r_2 \boldsymbol{e}_2 \cdot \boldsymbol{e}_i + r_3 \boldsymbol{e}_3 \cdot \boldsymbol{e}_i$$

$$\implies r_i = \boldsymbol{r} \cdot \boldsymbol{e}_i.$$

In Cartesian coordinates, the basis vectors $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$ are unit vectors pointing along the x, y and z-axes respectively. Furthermore, a conventional right-handed Cartesian coordinate system will obey $\hat{\boldsymbol{\imath}} \times \hat{\boldsymbol{\jmath}} = \hat{\boldsymbol{k}}$. Then, a general vector $\boldsymbol{A}$ in Cartesian coordinates can be written as

$$\boldsymbol{A} = (\boldsymbol{A} \cdot \hat{\boldsymbol{\imath}})\hat{\boldsymbol{\imath}} + (\boldsymbol{A} \cdot \hat{\boldsymbol{\jmath}})\hat{\boldsymbol{\jmath}} + (\boldsymbol{A} \cdot \hat{\boldsymbol{k}})\hat{\boldsymbol{k}} = A_x \hat{\boldsymbol{\imath}} + A_y \hat{\boldsymbol{\jmath}} + A_z \hat{\boldsymbol{k}},$$

where A_x, A_y and A_z are its components along the respective directions (Fig. 3.9). For the sake of convenience, the components of a vector are usually written as separate terms in brackets to save the need to mention the basis vectors.

$$\boldsymbol{A} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix}.$$

Multiplication by a Scalar

Multiplying a vector $\boldsymbol{A}$ by a scalar c is equivalent to multiplying each of its components by c, i.e.

$$c\boldsymbol{A} = c \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = \begin{pmatrix} cA_x \\ cA_y \\ cA_z \end{pmatrix}.$$

Addition of Vectors

The addition of two vectors can be performed by summing their corresponding components.

$$\boldsymbol{A}+\boldsymbol{B}=\begin{pmatrix}A_x\\A_y\\A_z\end{pmatrix}+\begin{pmatrix}B_x\\B_y\\B_z\end{pmatrix}=\begin{pmatrix}A_x+B_x\\A_y+B_y\\A_z+B_z\end{pmatrix}.$$

Dot Product

Since

$$\boldsymbol{A}=A_x\hat{\boldsymbol{\imath}}+A_y\hat{\boldsymbol{\jmath}}+A_z\hat{\boldsymbol{k}},$$

$$\boldsymbol{B}=B_x\hat{\boldsymbol{\imath}}+B_y\hat{\boldsymbol{\jmath}}+B_z\hat{\boldsymbol{k}},$$

and due to the facts that the dot product of two different basis vectors produces zero — as they are perpendicular — and that the dot product of identical basis vectors results in unity, we can apply the distributive property of the dot product to conclude that

$$\boldsymbol{A}\cdot\boldsymbol{B}=(A_x\hat{\boldsymbol{\imath}}+A_y\hat{\boldsymbol{\jmath}}+A_z\hat{\boldsymbol{k}})\cdot(B_x\hat{\boldsymbol{\imath}}+B_y\hat{\boldsymbol{\jmath}}+B_z\hat{\boldsymbol{k}})=A_xB_x+A_yB_y+A_zB_z$$

$$\boldsymbol{A}\cdot\boldsymbol{B}=\begin{pmatrix}A_x\\A_y\\A_z\end{pmatrix}\cdot\begin{pmatrix}B_x\\B_y\\B_z\end{pmatrix}=A_xB_x+A_yB_y+A_zB_Z. \tag{3.7}$$

A convenient way of remembering this result is that the dot product of two vectors is simply the sum of the product of their corresponding components. Substituting $\boldsymbol{B}=\boldsymbol{A}$ into the above equation, the magnitude of a vector is simply the square root of the sum of its squared components.

$$A=\sqrt{A_x^2+A_y^2+A_z^2}.$$

Cross Product

Expressing the vectors in terms of the basis vectors and applying the distributive property of the cross product as above, it can be shown that the cross product of $\boldsymbol{A}$ and $\boldsymbol{B}$ evaluates to

$$\boldsymbol{A}\times\boldsymbol{B}=\begin{pmatrix}A_x\\A_y\\A_z\end{pmatrix}\times\begin{pmatrix}B_x\\B_y\\B_z\end{pmatrix}=\begin{pmatrix}A_yB_z-A_zB_y\\A_zB_x-A_xB_z\\A_xB_y-A_yB_x\end{pmatrix}. \tag{3.8}$$

Note that in the derivation, we would have to use the facts that $\hat{\boldsymbol{\imath}}\times\hat{\boldsymbol{\jmath}}=\hat{\boldsymbol{k}}$, $\hat{\boldsymbol{k}}\times\hat{\boldsymbol{\imath}}=\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{\jmath}}\times\hat{\boldsymbol{k}}=\hat{\boldsymbol{\imath}}$. An easy way to remember the above result is as follows. To determine a particular component of the cross product $\boldsymbol{A} \times \boldsymbol{B}$, cover up that particular row. Then, multiply the component of $\boldsymbol{A}$ in the row below (loop to the top if necessary) with the diagonal component of $\boldsymbol{B}$ and subtract it by the product of the component of $\boldsymbol{A}$ two rows below, and the diagonal component of $\boldsymbol{B}$. For example, for the y-component

$$\begin{pmatrix} A_x \\ - \\ A_z \end{pmatrix} \times \begin{pmatrix} B_x \\ - \\ B_z \end{pmatrix},$$

the second row is covered. Then, we take A_z multiplied by the diagonal component B_x and subtract it by A_x multiplied by B_z, such that the y-component of the cross product is $A_z B_x - A_x B_z$.

3.2 Kinematic Quantities

Observational Frames of Reference

Firstly, kinematic quantities, such as velocity and acceleration, must be measured with respect to an observational frame of reference. This simply means that the motion is viewed from the perspective of a certain observer. Different observers can lead to vast differences in the quantities measured. Consider a system of three people labelled A, B and C respectively. From the perspective of C, A and B are running at speeds v_a and v_b ($v_b > v_a$) in the positive x-direction respectively. Then, it is intuitive from common experience that in the frame of A, B travels at a speed slower than v_b.

In light of the necessity of a reference frame, the lab frame is defined to be the frame where the laboratory, in which the "experiment" is conducted, is at rest. For most of this book, we shall omit mentioning "with respect to a frame of reference" for the sake of convenience and assume that all quantities are observed with respect to the lab frame, by default.

Coordinate Systems

A coordinate system is a construct we use to quantify positional measurements made in an observational reference frame. Note that a single observational reference frame can have multiple coordinate systems. The benefit of a coordinate system is that we can decompose spatial vectors into solely their components in the direction of the spatial coordinate axes. This is because the coordinate system implicitly implies direction, via its basis vectors. For example, (x, y, z) really means $x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}$. Common spatial coordinate systems include the Cartesian and polar coordinate systems which will both be introduced later.

Finally, location alone is not sufficient to define an event. After all, there is a need to know where and when something happens. Therefore, an event should also be described in terms of its time of occurrence. Furthermore, this should be a property of the frame of reference and independent of the spatial coordinate system.

Time

In the context of physics, the time of an event is the reading of the observer's clock[5] in that frame, when the event occurs. In classical mechanics, time is deemed as an invariant property which does not depend on the frame of reference. That is, if two observers synchronized their clocks at a certain instant and began to exhibit various forms of motion — which include traveling at a certain velocity and accelerating — the readings of their clocks will still be the same when they compare it at a later instant. Therefore, time unequivocally delineates the sequence of events that occur in all frames of reference.

That said, time is also relative in the sense that we usually mean the time elapsed between now and a previous instance when referring to time. Measuring elapsed time is akin to measuring length with a ruler — we take the difference between two measurements at separate points. For purposes of convenience however, we usually insinuate that the reading of the observational clock is set to zero when the experiment begins such that the elapsed time between now and the start is simply its current reading.

Displacement

The displacement of a particle, $\boldsymbol{s}$, is the change in the particle's position vector $\boldsymbol{r}$, from an initial state to a final state.

$$\boldsymbol{s} = \Delta \boldsymbol{r}.$$

To underscore the importance of displacement as a vector, consider a particle which has completed one full round of a circular track. In this case, the particle's position vector remains unchanged — implying that its displacement is the null vector. However, the distance travelled by the object is $2\pi R$

[5]Technically, the time of an event is the reading of a clock at the position of the event. However, the exact location of the clock does not matter for our current purposes as time is assumed to be universal in classical mechanics.

where R is the radius of the circle — a quantity that is non-zero. Evidently, displacement is starkly different from distance — the latter simply means the total ground covered during its motion and is a scalar.

Velocity

Velocity is defined as the rate of change of displacement and is also the rate of change of a particle's position vector (as $\boldsymbol{r} = \boldsymbol{s} + \boldsymbol{r}_0$ where the last term is the constant initial position vector that disappears upon differentiation).

$$\boldsymbol{v} = \frac{d\boldsymbol{s}}{dt} = \frac{d\boldsymbol{r}}{dt}.$$

Hence, the displacement of an object can be written as

$$\boldsymbol{s} = \int_0^t \boldsymbol{v} dt,$$

where we have set the reading of the clock to $t = 0$ at the start of the motion. If the object travels purely rectilinearly, the signed magnitude of displacement (i.e. the magnitude with a sign which denotes positive or negative) between times $t = t_0$ and $t = t_1$ is simply the positive area below the velocity-time graph between the two times (a negative velocity corresponds to a negative area).

To again differentiate between the velocity of a particle and its speed (which is a scalar), consider the case where it undergoes uniform circular motion. Note that speed is defined as the rate of change of distance covered. Though the particle's speed is constant in uniform circular motion, its velocity is varying as the direction of its velocity changes. Therefore, speed only reflects the magnitude of the velocity and not its direction.

Now, a common trap that many fall into is to hastily conclude that since $\frac{d\boldsymbol{r}}{dt}$ gives the velocity $\boldsymbol{v}$, $\frac{dr}{dt}$ should give the speed. This reasoning is fallacious as the latter gives the rate of change of the magnitude of the position vector which is the rate of change of the particle's distance from the origin dr and not the rate of change of distance $|d\boldsymbol{r}|$. Instead, the speed should be computed as $|\frac{d\boldsymbol{r}}{dt}|$. To emphasize the difference between speed and $\frac{dr}{dt}$, consider the following example.

Problem: A particle moves along the xy-plane at a constant velocity of $10\,\text{ms}^{-1}$ in the positive x-direction. When the particle is at $(3, 4)$, determine $\frac{d\boldsymbol{r}}{dt}$, its speed and $\frac{dr}{dt}$.

Let (x, y) denote the particle's instantaneous coordinates, so that

$$\frac{d\boldsymbol{r}}{dt} = \frac{d(x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}})}{dt} = \frac{dx}{dt}\hat{\boldsymbol{\imath}} + \frac{dy}{dt}\hat{\boldsymbol{\jmath}},$$

where the derivatives of $\hat{\boldsymbol{\imath}}$ and $\hat{\boldsymbol{\jmath}}$ have not been considered as they are constant.[6] Since $\frac{dx}{dt} = 10\text{ms}^{-1}$ and $\frac{dy}{dt} = 0$,

$$\frac{d\boldsymbol{r}}{dt} = 10\,\hat{\boldsymbol{\imath}}\,\text{ms}^{-1}.$$

The speed of the particle at this juncture is

$$\left|\frac{d\boldsymbol{r}}{dt}\right| = \sqrt{10^2 + 0} = 10\,\text{ms}^{-1}.$$

Now, the physical meaning of r is the particle's distance from the origin.

$$r^2 = x^2 + y^2$$
$$\implies 2r\frac{dr}{dt} = 2x\frac{dx}{dt} + 2y\frac{dy}{dt}.$$

Substituting $r = 5$ as $(x, y) = (3, 4)$, $\frac{dx}{dt} = 10$ and $\frac{dy}{dt} = 0$,

$$\frac{dr}{dt} = 6\,\text{ms}^{-1},$$

which is evidently an ineffective descriptor of the particle's motion.

Problem: Determine the condition for two particles with initial position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ and constant velocities $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ to collide.

The position vectors of the two particles after time t are $\boldsymbol{r}_1 + \boldsymbol{v}_1 t$ and $\boldsymbol{r}_2 + \boldsymbol{v}_2 t$ respectively. For them to coincide,

$$\boldsymbol{r}_1 + \boldsymbol{v}_1 t = \boldsymbol{r}_2 + \boldsymbol{v}_2 t$$
$$\implies \boldsymbol{r}_1 - \boldsymbol{r}_2 = (\boldsymbol{v}_2 - \boldsymbol{v}_1)t.$$

Since t must be a positive scalar (which excludes the possibility of the vectors being anti-parallel), the above implies that $\boldsymbol{v}_2 - \boldsymbol{v}_1$ is parallel to $\boldsymbol{r}_1 - \boldsymbol{r}_2$. That is, their unit vectors must be identical.

$$\frac{\boldsymbol{r}_1 - \boldsymbol{r}_2}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = \frac{\boldsymbol{v}_2 - \boldsymbol{v}_1}{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}.$$

[6] Note that not all basis vectors are constant — we shall see so when we encounter polar coordinates later.

Acceleration

Acceleration is defined as the rate of change of velocity,

$$\boldsymbol{a} = \frac{d\boldsymbol{v}}{dt}.$$

The velocity is thus related to the acceleration as

$$\int_{\boldsymbol{u}}^{\boldsymbol{v}} d\boldsymbol{v} = \int_0^t \boldsymbol{a} dt$$

$$\boldsymbol{v} = \boldsymbol{u} + \int_0^t \boldsymbol{a} dt,$$

where $\boldsymbol{u}$ is the initial velocity of the object. Similarly, if the object only accelerates along a straight line, the signed magnitude of the change in velocity between times $t = t_0$ and $t = t_1$ is simply the positive area below the acceleration-time graph between the two times.

Well, we could continue differentiating to get quantities such as the rate of change of acceleration (known as the jerk), but further derivatives are not physically meaningful. This is because, the state of a classical system of particles can be uniquely defined by their coordinates (which may be angular) and the first time-derivative of these coordinates. As long as an initial state is defined in terms of these quantities, all future states can be predicted via the physical laws. However, it is useful to define second time derivatives of these coordinates (which are akin to the accelerations) as the physical laws are often expressed in terms of these variables (and not further derivatives) — indirectly allowing us to determine the future states of a system. Ultimately, as the physical laws do not encode any information about the higher-order derivatives of a system, they are strictly derived from previously known quantities and are not illuminating.

Quantities in Terms of Scalars

Often, the dynamical laws generate vector equations which actually encapsulate three independent quantities along three independent spatial directions. Then, it is convenient to divide the vector equations into three components, relative to a coordinate system, to produce three separate scalar equations in most cases. For example $\boldsymbol{v}$ could become (v_x, v_y, v_z) in terms of Cartesian coordinates. Then the equations of motion, which describe how the motion of a system evolves over time and are derived from dynamical laws, can often be solved in a simpler fashion.

3.3 Constant Acceleration

We shall now derive the kinematic equations for the one-dimensional motion of a particle with an initial velocity $\boldsymbol{u}$, undergoing a constant acceleration $\boldsymbol{a}$ along the direction of $\boldsymbol{u}$. Since the problem is one-dimensional, we can simply consider the relevant components along this direction. The velocity, v is then

$$v = u + \int_0^t a dt = u + at. \tag{3.9}$$

The displacement is

$$s = \int_0^t v dt = ut + \frac{1}{2}at^2. \tag{3.10}$$

Interestingly, we note that the average velocity, $\bar{v}$ which is defined as the total displacement divided by the total time elapsed, is

$$\bar{v} = \frac{s}{\Delta t} = u + \frac{1}{2}at = \frac{v + u}{2}. \tag{3.11}$$

Hence, the average velocity between two instances is simply the average of the initial and final velocities. Lastly, there is another useful relationship. Multiplying Eq. (3.10) by $2a$,

$$2as = 2aut + a^2t^2.$$

Furthermore, we know that

$$v^2 = (u + at)^2 = u^2 + 2aut + a^2t^2.$$

Then,

$$v^2 = u^2 + 2as. \tag{3.12}$$

These four equations form the kinematics equations that describe the motion under a constant acceleration. They enable us to solve for relevant quantities in terms of one another.

Problem: Two cars are traveling towards each other at speeds u_1 and u_2 respectively. If their brakes apply a constant deceleration a, what must the minimum initial distance between the cars be so that they can stop before colliding into each other?

The final speeds of the car are zero. Applying Eq. (3.12) to the first car,

$$0 = u_1^2 - 2as_1.$$

The distance that the first car travels before stopping is

$$s_1 = \frac{u_1^2}{2a}.$$

Similarly for the second car, $s_2 = \frac{u_2^2}{2a}$. The minimum initial distance is then the sum of s_1 and s_2:

$$s_1 + s_2 = \frac{u_1^2 + u_2^2}{2a}.$$

Projectile Motion

Projectile motion refers to the motion of a body under the influence of a gravitational force. An object under free-fall undergoes an approximately constant acceleration g downwards in the regime that we are considering. We shall not inquire about the force that causes this acceleration for now and will only deal with the trajectory of the particle. This is a general trend in mechanics. One first obtains the equations of motion of a system from the dynamical laws, after which, the problem is strictly a kinematic one.

As the acceleration of a projectile is constant, the motion of the projectile is strictly confined to a plane.[7] Hence, only two spatial coordinates are required to define the location of the projectile. An important property is that the motions of the body along any two perpendicular directions are independent as its acceleration is constant. Hence, this two-dimensional motion can be effectively divided into two separate one-directional motions whose kinematics equations can each be solved for. Along the vertical y-direction, the evolution of the particle's y-coordinate is akin to that of a particle undergoing a constant acceleration in a one-dimensional motion. Along the horizontal x-direction, the evolution of the particle's x-coordinate is equivalent to that of a particle travelling at a constant velocity in one dimension.

Let $x(t)$ and $y(t)$ be the horizontal and vertical coordinates of the projectile and let u_x and u_y be its horizontal and vertical velocities at time $t = 0$. We take the positive direction of the y-axis to be vertically upwards. Then,

[7] The reader should try to prove this. Hint: define the origin at the initial position of the particle and show that the position vector of the particle thereafter is always perpendicular to a certain vector (to find this vector, apply a certain vector operation to $\boldsymbol{u}$ and $\boldsymbol{a}$).

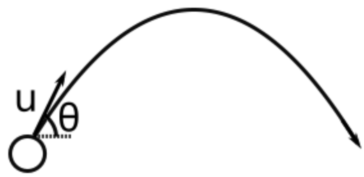


Figure 3.10: Trajectory of a projectile

if we let x_0 and y_0 be the horizontal and vertical coordinates of the projectile at time $t = 0$, the kinematics equations yield

$$x = x_0 + s_x = x_0 + u_x t, \tag{3.13}$$

$$y = y_0 + s_y = y_0 + u_y t - \frac{1}{2} g t^2. \tag{3.14}$$

There are certain interesting properties regarding the trajectory of a projectile.

Property 1: The general shape of the trajectory is a parabola with its vertex at $(x_0 + \frac{u_x u_y}{g}, y_0 + \frac{u_y^2}{2g})$.

Proof: t can first be expressed in terms of x:

$$t = \frac{x - x_0}{u_x}.$$

Substituting this expression for t into $y(t)$,

$$y = y_0 + \frac{u_y}{u_x}(x - x_0) - \frac{g}{2u_x^2}(x - x_0)^2. \tag{3.15}$$

Completing the square,

$$y = y_0 + \frac{u_y^2}{2g} - \frac{g}{2u_x^2}\left(x - x_0 - \frac{u_x u_y}{g}\right)^2,$$

which is the equation of an inverted parabola. Of course, the particle will probably not traverse the entire parabola due to impediments such as the ground. However, this parabolic equation is valid for the regime in which the particle is still under free-fall. It can be seen that the vertex of the parabola is at $(x_0 + \frac{u_x u_y}{g}, y_0 + \frac{u_y^2}{2g})$. Hence, the maximum y-coordinate that the projectile can reach is $y_0 + \frac{u_y^2}{2g}$, assuming that it has not passed this point yet. Often, the more edifying form of the above expression is obtained by expressing the initial velocity in polar coordinates. u is the initial speed of the projectile while θ is the angle subtended by the initial velocity and the positive x-axis (Fig. 3.10). Substituting $u_y = u\sin\theta$ and $u_x = u\cos\theta$ into Eq. (3.15),

$$y = y_0 + \tan\theta(x - x_0) - \frac{g\sec^2\theta}{2u^2}(x - x_0)^2. \tag{3.16}$$

The usefulness of this expression is evident in problems where the angle θ is a variable. Lastly, since the parabola is defined by the velocity at a single state along the path, if the projectile starts at a different position along the parabola and possesses the corresponding velocity, the resultant parabola will be the same as before.

Property 2: Projectile motion is reversible. That is, if the direction of the velocity of the projectile at a certain juncture is reversed, the trajectory will still take the form of the same parabola.

Proof: Replacing θ with $\pi + \theta$ in Eq. (3.16) does not modify the trajectory equation.

Property 3: Given a fixed vertical displacement s, the time elapsed during projectile motion is independent of its horizontal motion. In fact, the time elapsed can be represented in terms of s and h, the maximum height above the initial vertical level. We still adopt the same definition for h in the case where the particle has already crossed the point of maximum height.

Proof:

$$s = u_y t - \frac{1}{2}gt^2.$$

Furthermore, the maximum height, h, can be computed by setting $v = 0$ in Eq. (3.12):

$$0 = u_y^2 - 2gh$$

$$h = \frac{u_y^2}{2g}.$$

We wish to express u_y in terms of h.

$$u_y = \pm\sqrt{2gh}.$$

The positive value corresponds to situations where the projectile has yet to reach the maximum height at $t = 0$ (i.e. it is traveling vertically upwards).

The negative value corresponds to the converse. Then,

$$s = \pm\sqrt{2gh}t - \frac{1}{2}gt^2$$

$$t^2 \mp 2\sqrt{\frac{2h}{g}}t + \frac{2s}{g} = 0$$

$$t = \pm\sqrt{\frac{2h}{g}} \pm \sqrt{\frac{2(h-s)}{g}}.$$

If $u_y > 0$ and $s > 0$ (i.e. the projectile is initially traveling upwards and ends its trajectory at a higher position),

$$t = \sqrt{\frac{2h}{g}} \pm \sqrt{\frac{2(h-s)}{g}}.$$

In the case where $u_y > 0$ and $s < 0$ (i.e. the projectile is initially traveling upwards and ends its trajectory at a lower position), there is only one physically reasonable solution:

$$t = \sqrt{\frac{2h}{g}} + \sqrt{\frac{2(h-s)}{g}}.$$

Lastly, when $u_y < 0$, the only physically possible range for s is $s < 0$ (i.e. the projectile is initially traveling downwards and ends up at a lower position). Then,

$$t = -\sqrt{\frac{2h}{g}} + \sqrt{\frac{2(h-s)}{g}}.$$

Hence, given a starting and ending point on the parabola of a projectile's trajectory, the time elapsed between these two points can be expressed in terms of the vertical displacement s and the vertical distance between the starting point and the vertex of the parabola, h. Note that there was completely no mention of the horizontal coordinates. We can then analyze how t varies with h.

In the positive case of the first situation and the second situation, the time t is evidently smaller when the maximum height h is smaller. In the negative case of the first situation and the last situation, the time t is smaller when the maximum height h is larger. This can be seen by considering the derivatives. For example, in the negative case of the first situation,

$$t = \sqrt{\frac{2h}{g}} - \sqrt{\frac{2(h-s)}{g}}.$$

Differentiating this with respect to h,

$$\frac{dt}{dh} = \frac{1}{\sqrt{2gh}} - \frac{1}{\sqrt{2gh - 2gs}} < 0.$$

This also shows that $\frac{dt}{dh} < 0$ in the third equation, which is negative of the equation we have just differentiated but with $s < 0$. In conclusion, if the particle will cross the point of maximum height, t decreases with a smaller height h. This is intuitive as it is expeditious for the particle to have a smaller velocity upwards to swiftly reach the maximum point before falling down. Otherwise if it has already passed or will not reach the point of maximum height, t decreases with a larger height h. This is also intuitive as it implies that the particle has a greater initial velocity and will reach the required point in a shorter time interval.

Problem: Tom stands on top of a tower that is of a height s above the horizontal ground. He then tosses two darts, A and B, simultaneously at two corresponding targets A and B that are at horizontal distances x_1 and x_2 relative to the tower, respectively, with $x_1 < x_2$. These two targets are stationed on the ground. If the maximum heights attained by darts A and B are y_1 and y_2, relative to the ground, respectively, with $y_1 > y_2$, which dart hits their corresponding target first?

Since the darts cross their respective points of maximum height and traverse the same vertical displacement, dart B takes a shorter time to reach its target, despite the larger horizontal distance that it has to cover, as it attains a lower maximum height.

3.3.1 *Projectile Motion with Drag*

When an object travels in a fluid medium such as air, it experiences a drag force of the form

$$\boldsymbol{F}_{drag} = -bv\hat{\boldsymbol{v}} - cv^2\hat{\boldsymbol{v}} = \boldsymbol{F}_{lin} + \boldsymbol{F}_{quad}$$

which is opposite in direction to $\boldsymbol{v}$, the object's velocity relative to the medium. b and c are constants that depend on various parameters. There are two components, one with a linear dependence on the speed and one with a quadratic dependence on the speed. These have different physical origins. The linear term is attributed to the viscosity of the fluid medium while the quadratic term is due to the collisions of the object with neighboring fluid molecules — accelerating them in the process. Most of the time, one component will be dominant over the other.

Generally, when the speed of the object is small and the fluid medium is highly viscous, the linear term dominates. However, when the speed of the object is large, the quadratic term tends to dominate. Hence, in most problems involving drag, only one component of the drag force is considered. Unfortunately, in most realistic situations of projectile motion with drag, the quadratic term dominates. The trajectory equation cannot be analytically solved from the nasty coupled differential equation. However, we can still make a few qualitative arguments.

Qualitative Properties of Projectile Motion with Drag

The following summarizes several important properties:

- The maximum height attained by the projectile in a situation with drag is lower than that in an ideal situation without drag. This is due to the fact that the vertical component of the drag force is directed downwards while the projectile is traveling upwards, causing the projectile to decelerate vertically at a magnitude greater than g.
- The trajectory is now asymmetric about the vertical line crossing the point of maximum height. If the projectile is launched from a flat ground, the time taken for the projectile to reach the maximum height is less than the time required for the projectile to land back onto the ground from the maximum height. This is due to the fact that the drag force is directed downwards as the projectile travels upwards (reinforcing its deceleration due to g) but is directed upwards as the projectile falls downwards (opposing the acceleration due to g).
- The horizontal range of the projectile is generally smaller. Unfortunately, there is no rigorous justification of this. That said, consider the following factors. Firstly, the horizontal speed decreases with time. Next, the total time of flight is likely to be less than that of projectile motion without drag. Though, the time required for the ascent is less while that for descent is more than the corresponding cases without drag. The additional decrease in time is likely to be more significant than the additional increase in time as the magnitude of the drag force during the ascent (when the speed of the object is large) is likely larger than that during the descent (when the speed of the object is small). All-in-all, though these factors are not

perfectly rigorous arguments, they provide a rough explanation for the decrease in range observed in numerical computations.

3.4 Polar Coordinates

So far, we have been considering a Cartesian coordinate system. In this section, a different coordinate system, known as the polar coordinate system will be introduced. In polar coordinates, we can describe each point in two-dimensional space with a magnitude r corresponding to the distance of that point from the origin and an angle θ subtended by the position vector $\boldsymbol{r}$ of that point and a fixed axis which is usually the x-axis. Referring to Fig. 3.11 below, the basis vectors used to describe a spatial position are $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\theta}}$, which are the radial and tangential unit vectors respectively.

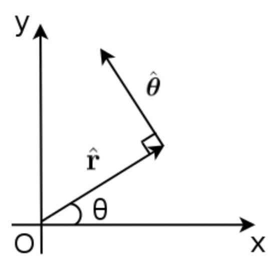


Figure 3.11: Polar coordinates

The advantage of this coordinate system is that spatial positions are directly described by their distances to the origin, r, as opposed to the indirect distance $\sqrt{x^2+y^2}$ in the case of a Cartesian coordinate system. Polar coordinates are especially convenient when we know how the distance r or angle θ varies. However, this convenience is at the expense of the basis vectors $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\theta}}$ changing as θ varies (but not r).

The position vector of a particle at a distance r away from the origin is

$$\boldsymbol{r} = r\hat{\boldsymbol{r}},$$

while the velocity of the particle is the rate of change of its position vector, i.e.

$$\boldsymbol{v} = \frac{d\boldsymbol{r}}{dt} = \dot{r}\hat{\boldsymbol{r}} + r\dot{\hat{\boldsymbol{r}}},$$

where a dot refers to a total time derivative. Note that there is a need to consider the derivative of the radial unit vector as it is no longer a constant.

We can evaluate the derivatives of the basis vectors by first expressing them in Cartesian coordinates.

$$\hat{\boldsymbol{r}} = \cos\theta\hat{\boldsymbol{i}} + \sin\theta\hat{\boldsymbol{j}},$$
$$\hat{\boldsymbol{\theta}} = -\sin\theta\hat{\boldsymbol{i}} + \cos\theta\hat{\boldsymbol{j}}.$$

A quick derivation of the above is to observe that $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{\theta}}$ are $\hat{\boldsymbol{i}}$ and $\hat{\boldsymbol{j}}$ rotated anti-clockwise by θ. Consequently, we can apply the rotation matrix to $(\hat{\boldsymbol{i}} \quad \hat{\boldsymbol{j}})$.

$$\begin{aligned}(\hat{\boldsymbol{r}} \quad \hat{\boldsymbol{\theta}}) &= \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}(\hat{\boldsymbol{i}} \quad \hat{\boldsymbol{j}}) = \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & 0 \\ 0 & 1\end{pmatrix} \\ &= \begin{pmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{pmatrix},\end{aligned}$$

which just returns the rotation matrix[8]

$$\implies \hat{\boldsymbol{r}} = \begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix} \quad \hat{\boldsymbol{\theta}} = \begin{pmatrix}-\sin\theta \\ \cos\theta\end{pmatrix}.$$

Differentiating the polar basis vectors with respect to time,

$$\dot{\hat{\boldsymbol{r}}} = -\sin\theta\dot{\theta}\hat{\boldsymbol{i}} + \cos\theta\dot{\theta}\hat{\boldsymbol{j}} = \dot{\theta}\hat{\boldsymbol{\theta}},$$
$$\dot{\hat{\boldsymbol{\theta}}} = -\cos\theta\dot{\theta}\hat{\boldsymbol{i}} - \sin\theta\dot{\theta}\hat{\boldsymbol{j}} = -\dot{\theta}\hat{\boldsymbol{r}}.$$

Hence,

$$\boldsymbol{v} = \dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}.$$

How can we understand these two components? The first term can be seen as the radial velocity

$$\boldsymbol{v}_r = \dot{r}\hat{\boldsymbol{r}}, \tag{3.17}$$

whose magnitude describes the rate of change of the "radius" of the particle to the origin. The second term is the tangential velocity

$$\boldsymbol{v}_\theta = r\dot{\theta}\hat{\boldsymbol{\theta}}, \tag{3.18}$$

[8] A slick derivation of the rotation matrix is as follows: define (x, y) and (x', y') as the coordinates of a point before and after an anti-clockwise rotation of it over angle θ. Next, define new complex variables $\eta = x + iy$ and $\eta' = x' + iy'$ which represent these coordinates in the complex plane. Since an anti-clockwise rotation of a complex number in the complex plane by angle θ is performed by multiplying $e^{i\theta}$ to it, $\eta' = e^{i\theta}\eta = (\cos\theta + i\sin\theta)(x + iy) = \cos\theta x - \sin\theta y + i(\sin\theta x + \cos\theta y)$. x' and y' can then be retrieved from η' by taking its real and complex components. $x' = \cos\theta x - \sin\theta y$ while $y' = \sin\theta x + \cos\theta y$.

whose magnitude indirectly describes the rate of change of θ and indirectly quantifies how fast the position vector of the particle is "rotating" about the origin. $\dot{\theta}$ is known as the angular velocity. Next, the acceleration is the rate of change of velocity.

$$\begin{aligned}\boldsymbol{a} &= \frac{d(\dot{r}\hat{\boldsymbol{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}})}{dt} \\ &= \ddot{r}\hat{\boldsymbol{r}} + \dot{r}\dot{\hat{\boldsymbol{r}}} + \dot{r}\dot{\theta}\hat{\boldsymbol{\theta}} + r\ddot{\theta}\hat{\boldsymbol{\theta}} + r\dot{\theta}\dot{\hat{\boldsymbol{\theta}}} \\ &= (\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{r}} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{\theta}}.\end{aligned}$$

Splitting the acceleration into radial and tangential components,

$$\boldsymbol{a}_r = \left(\ddot{r} - r\dot{\theta}^2\right)\hat{\boldsymbol{r}}, \tag{3.19}$$

$$\boldsymbol{a}_\theta = \left(r\ddot{\theta} + 2\dot{r}\dot{\theta}\right)\hat{\boldsymbol{\theta}}. \tag{3.20}$$

$\ddot{\theta}$ is known as the angular acceleration. Perhaps, two facts that we would need to get used to are that the radial acceleration is not directly equal to $\ddot{r}$ and that the tangential acceleration is not directly equal to $r\ddot{\theta}$. These are due to the fact that the direction of the velocity vector also varies with respect to time. We can understand each term in the two equations above in greater detail.

$\ddot{r}$ is the rate of change of the radial speed, $\dot{r}$. Next, $-r\dot{\theta}^2$ is known as the centripetal acceleration and corresponds to the instantaneous radial acceleration required for the particle to instantaneously travel in an arc of constant radius about the origin. The negative sign indicates that this acceleration is directed radially inwards.

$r\ddot{\theta}$ indirectly describes the angular acceleration $\ddot{\theta}$. It is the component of the tangential acceleration that causes the angular velocity $\dot{\theta}$ of the particle to change. $2\dot{r}\dot{\theta}$ is known as the Coriolis acceleration and unfortunately does not have an intuitive physical meaning.

Problem: A particle moves at a constant $\dot{\theta} = \omega$ and with $r = r_0 e^{\alpha t}$. Show that for some values of α, $a_r = 0$.

$$a_r = \ddot{r} - r\dot{\theta}^2 = \alpha^2 r_0 e^{\alpha t} - \omega^2 r_0 e^{\alpha t}$$

$$\implies a_r = 0 \quad \text{when} \quad \alpha = \pm\omega.$$

It may be surprising to many that a_r can be zero even though the radial velocity $\dot{r} = \alpha r_0 e^{\alpha t}$ is increasing with time — an apparent paradox. This underscores the ramifications of the changing directions of the basis vectors. This apparent paradox arises from the misconception that a_r solely contributes to $\ddot{r}$, which is not true as we have neglected the centripetal acceleration. In other words

$$\dot{r} \neq \int a_r(t)dt,$$

as $a_r = \ddot{r} - r\dot{\theta}^2 \neq \ddot{r}$.

Problem: A particle moves at an angular velocity described by $\dot{\theta} = \alpha t$ for some constant angular acceleration α. If its radius from the origin evolves with respect to time according to the equation $r^2 = \frac{k}{\dot{\theta}}$ for some constant k, show that the tangential acceleration a_θ is zero.

$$a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta} = \sqrt{\frac{k}{\alpha t}} \cdot \alpha - 2 \cdot \frac{1}{2}\sqrt{\frac{k}{\alpha t^3}} \cdot \alpha t = 0.$$

This may also be surprising in the sense that the angular velocity of the particle is increasing though there is no tangential acceleration. However, this is due to the fact that

$$a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta} \neq r\ddot{\theta}.$$

The tangential acceleration does not solely contribute to the angular acceleration. The Coriolis acceleration must also be accounted for. In fact, we can prove something more general; if the quantity $r^2\dot{\theta}$ is a constant, $a_\theta = 0$.

$$a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta} = \frac{r^2\ddot{\theta} + 2r\dot{r}\dot{\theta}}{r} = \frac{\frac{dr^2\dot{\theta}}{dt}}{r} = 0.$$

Referring to the next problem, polar coordinates are especially useful when certain kinematic quantities, such as velocity or acceleration, are always radial or tangential.

Problem: A rabbit is running at a constant velocity v rectilinearly in the x-direction. A tiger chases the rabbit at instantaneous velocity u ($u > v$) of constant magnitude and its direction is always pointing from the instantaneous position of the tiger to the instantaneous position of the rabbit. What is the time taken for the tiger to catch the rabbit? The initial distance between them is l and the initial position vector of the rabbit relative to the tiger makes an angle θ_0 with the positive x-axis (the tiger and the rabbit lie on the same plane).

Well, a natural step to take would be to consider the frame of the tiger (we will use the same x and y-axes as the lab frame). Even though we have

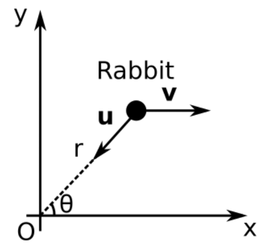


Figure 3.12: Frame of tiger

not introduced the principle of Galilean relativity yet, it should be intuitive that if two objects A and B are traveling at velocities $\boldsymbol{v}_A$ and $\boldsymbol{v}_B$ in the lab frame, the velocity of B in the frame of A is given by $\boldsymbol{v}_B - \boldsymbol{v}_A$. Hence the velocity of the rabbit in the frame of the tiger is $\boldsymbol{v} - \boldsymbol{u}$, as depicted in the figure above. Well, if we solely consider Cartesian coordinates, we can write down the following integrals by letting τ be the time the tiger takes to catch the rabbit. In order for them to meet at the same y-coordinate,

$$\dot{y} = -u \sin\theta,$$

$$-\int_0^\tau u \sin\theta \, dt = -l \sin\theta_0. \tag{3.21}$$

Similarly, for them to coincide at the same x-coordinate,

$$\dot{x} = v - u\cos\theta,$$

$$\int_0^\tau (v - u\cos\theta)\, dt = -l\cos\theta_0. \tag{3.22}$$

Evidently, we have reached a impasse as these integrals cannot be evaluated explicitly. The trick here is to consider the situation in both Cartesian and polar coordinates. Consider the velocity of the rabbit along the radial direction in the frame of the tiger,

$$\dot{r} = v\cos\theta - u$$

$$\int_0^\tau (v\cos\theta - u)\, dt = -l. \tag{3.23}$$

Now, we can see a way out of this mess. By multiplying Eq. (3.22) by v and adding it to Eq. (3.23) multiplied by u,

$$\int_0^\tau (v^2 - u^2)\, dt = -l(v\cos\theta_0 + u).$$

Since the integrand is constant, the integral on the left-hand side can be evaluated trivially. Then,

$$\tau = \frac{l(v\cos\theta_0 + u)}{u^2 - v^2}.$$

Let's retrace how we solved this problem. Usually, to ensure that two point particles are coincident, we must check that both their x and y-coordinates are the same. However, applying this method to this situation would only generate a pair of integrals that are impossibly difficult to solve for. Instead, we can ensure that the x and r coordinates of those particles coincide. This is due to the fact that the r coordinate is not completely dependent on the x-coordinate and is also a function of the y coordinate. As such, if the r and x coordinates of two objects are identical, their y coordinates must also be the same. Hence, we have witnessed how some problems are naturally suited to be expressed in polar coordinates.

3.4.1 *Uniform Circular Motion*

Uniform circular motion refers to the motion of an object in a circle of constant radius at a constant tangential speed and thus angular velocity. Note that though the object's tangential speed is constant, it is still undergoing acceleration as the direction of its velocity is constantly changing. Expressing these conditions in polar coordinates ($\dot{r} = 0$, $\ddot{r} = 0$ and $\ddot{\theta} = 0$), we obtain

$$a_r = \ddot{r} - r\dot{\theta}^2 = -r\dot{\theta}^2,$$

$$a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta} = 0.$$

We find that the object must undergo an acceleration radially inwards which we term as the centripetal acceleration. $-r\dot{\theta}^2$ is the amount of radial acceleration necessary for the object to remain traveling in a circle of radius r when it has an angular velocity $\dot{\theta}$. We shall also present a geometric proof of the required centripetal acceleration as follows.

Let the origin be at the center of rotation. Then, consider two points on the circular path that are separated by a small angular distance $d\theta$ and consider the position and velocity vectors at both points. The change in the position vector is simply the vector obtained from joining the heads of the position vectors — pointing from the object's initial position to its final position. A similar statement can be made about the change in the velocity vector. Then, if we group these vectors into triangles as shown on the right

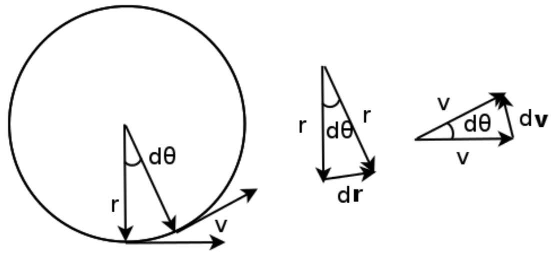


Figure 3.13: Uniform circular motion

of Fig. 3.13 above, we see that they are in fact similar (SAS). Thus,

$$\frac{|d\boldsymbol{v}|}{|d\boldsymbol{r}|} = \frac{v}{r} \implies \left|\frac{d\boldsymbol{v}}{dt}\right| = \frac{v}{r}\left|\frac{d\boldsymbol{r}}{dt}\right|.$$

Denoting $|\frac{d\boldsymbol{v}}{dt}|$ as a_c and because $|\frac{d\boldsymbol{r}}{dt}| = v$,

$$a_c = \frac{v^2}{r} = r\omega^2,$$

where we have used ω to denote the particle's angular velocity. Note that $|d\boldsymbol{r}|$ and $|d\boldsymbol{v}|$ refer to the magnitude of the vectorial change in r and v. In the limit where $d\theta \to 0$, the vector $d\boldsymbol{v}$ also points radially inwards. Hence, it is evident that there must be a centripetal acceleration for uniform circular motion.

3.4.2 *Circular Motion with Tangential Acceleration*

Now, consider the case in which the particle's tangential velocity is allowed to increase. However, it must still travel in a circle of constant radius r. Then, by setting $\dot{r} = 0$ and $\ddot{r} = 0$, its kinematics equations in polar coordinates become

$$a_r = -r\dot{\theta}^2,$$
$$a_\theta = r\ddot{\theta}.$$

This means that there must be an instantaneous centripetal acceleration $-r\dot{\theta}^2$, where $\dot{\theta}$ is the instantaneous angular velocity, for the particle to remain along the circular path. However, if a tangential acceleration is present, there will be an angular acceleration, causing the instantaneous angular velocity at the next instant to be different from that at the current instant. This means that the centripetal acceleration has to constantly change in magnitude to adapt to the changes in the angular velocity. In summary, the instantaneous angular velocity $\dot{\theta}$ may now be a function of time but the centripetal acceleration must always be $-r\dot{\theta}^2$ for the particle to remain in a circular path of radius r.

3.5 Kinematics of a Rigid Body

The previous sections analyzed the motion of a single, discrete particle. In this section, we shall consider the motion of a special system of particles. A macroscopic body is made up of myriad particles which each have three translational degrees of freedom. If it consists of n particles, a total of $3n$ generalized coordinates (translational or angular) is required to define a unique state of the system. However, this immense complexity is greatly simplified in the case of a rigid body.

A rigid body is a body whose particles maintain a constant distance relative to each other. This rigid body constraint is merely an idealization. Consider the following situation. When one end of a moving object is suddenly blocked by an impregnable wall, the other end of the object cannot instantaneously know that it must stop too, as there is a limit to the speed of "signals" in the material which is by definition, the speed of sound in that medium. Then, at the next instant, the particles that constitute the body will definitely be closer together, violating the criterion of a rigid body. Despite such impracticalities, the notion of a rigid body is still a rather convenient approximation.

Chasles's theorem states that the most general rigid body motion can be represented in terms of a translation of an arbitrary point P on the body and a rotation of the entire body about P. This is a result of the rigid body constraint and drastically reduces the number of coordinates[9] required to describe a state of a rigid body from $3n$ to 6. We will not prove this theorem but the result should be intuitive and believable.

3.5.1 *Angular Speed and Velocity*

Supposing that we choose a particular particle P in applying Chasles's theorem, consider the frame of P and define the origin at its location. Since the motion of the rigid body in P's frame is a rotation, define the xy-plane to be the instantaneous plane of rotation.

In cylindrical coordinates, every point on the body can be ascribed a coordinate θ, the angle between the projection of its position vector on the

[9]Usually, three coordinates describe the position of an arbitrary point P and the other three define the orientation of the body about P.

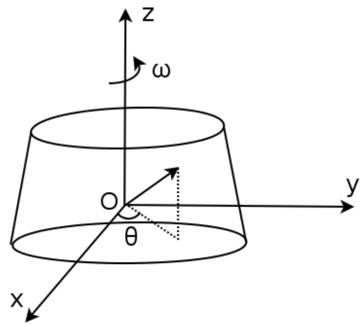


Figure 3.14: Rotation of a rigid body

xy-plane and the x-axis (Fig. 3.14). By definition of a rotation, if the θ coordinate of one off-axis particle increases by $d\theta$ during a time interval dt, the θ coordinate of every off-axis particle on the rigid body must also increase by $d\theta$. Note that we are considering infinitesimal rotations as the plane of rotation may change over time.

It can be seen that the quantities $d\theta$ and $\frac{d\theta}{dt}$ are representative of the entire rigid body. $\omega = \frac{d\theta}{dt}$ is the angular speed of the rigid body — it is the rate of change of the azimuthal angular coordinate of all points on the rigid body, relative to an arbitrary point. Now, you may think that ω is associated with a certain point (such as the point P that was chosen) or axis (such as the z-axis). However, this is not the case.

In the frame[10] of a different particle Q, the body still undergoes a rotation at the same angular speed ω in the same direction (clockwise or anticlockwise). This is evident from the fact that if the velocity of Q in the original frame is $\boldsymbol{v}$, the velocity of the particle P in the frame of Q is $-\boldsymbol{v}$. Since the distance between P and Q is the same in both frames, the angular speeds are identical in both frames. One can also easily check that the directions of rotation are also identical. Therefore, the angular speed is truly an intrinsic characteristic of the entire rigid body.

Finally, the angular speed of a rigid body, ω, is most importantly invariant (i.e. not changing) across frames that are non-rotating with respect to each other as it describes the relative motion of particles.

Next, the angular velocity vector of a rigid body, $\boldsymbol{\omega}$, is defined to be a vector perpendicular to the plane of rotation. Its magnitude reflects the angular speed ω. Notice that there are two possible directions for $\boldsymbol{\omega}$. By convention, the direction of $\boldsymbol{\omega}$ is defined by the right-hand grip rule. If you

[10]Note that the frame we choose is not rotating with respect to the frame of P.

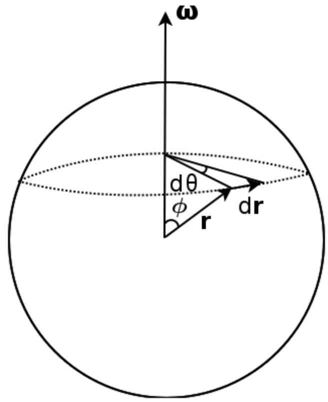


Figure 3.15: Angular velocity of a rigid body

point your right thumb in the direction of $\boldsymbol{\omega}$, your other fingers will curl in the direction of rotation (clockwise or anti-clockwise). For example, an angular velocity in the positive z-direction would imply an anti-clockwise rotation in the x-y plane. Finally, similar to the angular speed, $\boldsymbol{\omega}$ is a characteristic of the entire rigid body that is invariant across non-rotating frames.

With this definition, there is an elegant way of expressing the velocity of every point on the rigid body in the frame of an arbitrary particle P. Referring to the previous analysis, the angular velocity $\boldsymbol{\omega}$ is in the z-direction.

Now, consider a particle that is instantaneously at a position vector $\boldsymbol{r}$ from the reference particle. $\boldsymbol{r}$ makes an angle ϕ with the angular velocity vector $\boldsymbol{\omega}$ (Fig. 3.15). Then, after a time dt, the particle travels an angle $d\theta = \omega dt$ along a circular path of radius $r \sin\phi$. Hence, the magnitude of the infinitesimal displacement is

$$|d\boldsymbol{r}| = r\sin\phi d\theta,$$

and is directed azimuthally. The magnitude of this displacement, in combination with its direction, can be expressed cogently in terms of vectors as

$$d\boldsymbol{r} = \boldsymbol{\omega} \times \boldsymbol{r} dt$$

$$\frac{d\boldsymbol{r}}{dt} = \boldsymbol{\omega} \times \boldsymbol{r}. \tag{3.24}$$

In fact, $\boldsymbol{r}$ could be any vector in general that is rotating at an angular velocity $\boldsymbol{\omega}$ and the above derivation would still hold. Moving back to the main topic, $\frac{d\boldsymbol{r}}{dt}$ is the velocity of a particle in the frame of particle P. Suppose that particle P has a velocity $\boldsymbol{v}_{ref}$ in the lab frame, the velocity of a point on the rigid body, with respect to the lab frame, that is at a position vector $\boldsymbol{r}$

relative to P is

$$\boldsymbol{v} = \boldsymbol{v}_{ref} + \boldsymbol{\omega} \times \boldsymbol{r}. \tag{3.25}$$

This is an extremely important equation that relates the velocity of an arbitrary point on a rigid body to that of a reference point. Though the reference point can be taken to be any point on the rigid body, it is usually defined to be located at the center of mass or some fixed point in the lab frame, in practice. This is because the dynamical laws only enlighten us on the acceleration, and thus velocity, of the center of mass and not any other point on the rigid body. Therefore, the evolution of $\boldsymbol{v}_{ref}$ is impossible to determine directly for other non-fixed points that are not the center of mass.

In most cases, we will be dealing with objects moving and rotating in a single plane. It is then ideal to express the above equation in terms of two-dimensional polar coordinates attached to the reference point. $\boldsymbol{r}$ is then the position vector of a particle of concern. Since $\boldsymbol{\omega}$ is defined to be positive along the positive z-direction, which is perpendicular to the plane of motion, $\boldsymbol{\omega} \times \boldsymbol{r} = r\omega\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}$ is the tangential unit vector.

$$\boldsymbol{v} - \boldsymbol{v}_{ref} = r\omega\hat{\boldsymbol{\theta}}.$$

Define v_r and $v_{ref,\,r}$ as the respective speeds of the particle of concern and the reference particle in the lab frame, along the radial direction. Similarly, let v_t and $v_{ref,\,t}$ be those in the tangential direction respectively,

$$v_r - v_{ref,\,r} = 0,$$

$$v_t - v_{ref,\,t} = r\omega.$$

Therefore, the components of the velocities of two particles on a rigid body, along the line joining them, must be identical. This is rather intuitive: suppose that this were not the case, then the distance between them will change in the next instance. Furthermore, the relative tangential velocities must correctly reflect the angular speed. Notice that we could have also obtained these equations by substituting $\dot{r} = 0$ and $\ddot{r} = 0$ in the kinematics equation in polar coordinates.

Problem: If the angular velocity vector of rod AB is $\boldsymbol{\omega}_{AB} = -\omega\hat{\boldsymbol{k}}$, find the velocity of the pinned connection C in Fig. 3.16. All rods are assumed to be rigid. Note that all connections are pinned — implying that the connected components can rotate with respect to each other. The x and y-axes are positive rightwards and upwards.

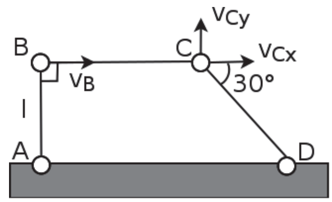


Figure 3.16: Connected rods

Let us first solve this problem by analyzing the velocities of the connection points. Points A and D are fixed to the ground and thus have zero velocity. The velocity of B is simply

$$v_B = l\omega,$$

in the positive x-direction. Now, consider the components of the velocity of point C parallel and perpendicular to rod BC, which we conveniently label as v_{Cx} and v_{Cy}. As two points on a rigid body cannot have a relative velocity along the line connecting them,

$$v_{Cx} = v_B = l\omega$$

for rod BC to remain rigid. Lastly, we impose this condition on rod CD. Since D is stationary, the component of point C's velocity along CD must be zero. That is,

$$v_{Cy} \cos 60^\circ = v_{Cx} \cos 30^\circ$$
$$v_{Cy} = \sqrt{3} l\omega.$$

Therefore,

$$\boldsymbol{v}_C = l\omega\hat{\boldsymbol{\imath}} + \sqrt{3} l\omega\hat{\boldsymbol{\jmath}}.$$

Instantaneous Center of Rotation

Equation (3.25) can be further simplified if we choose the reference point to be a point known as the instantaneous center of rotation (ICoR). The instantaneous center of rotation is defined as the point that is fixed to the frame of the rigid body — but not necessarily lying on the body — that has zero velocity in the lab frame at that particular instant. If the ICoR is chosen to be the reference point, the velocity of a point on the body is described by

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$$

as $\boldsymbol{v}_{ref} = \mathbf{0}$. Hence, all points on the rigid body are seemingly rotating about the ICoR at this particular instant. Perhaps, the concept of a point that is fixed to the frame of the rigid body but not on the body itself is quite confusing. Consider attaching an infinite plane onto the body like a giant sheet of paper. This paper is fixed to the rigid body (i.e. it rotates with the body) and represents space as observed in the frame of the rigid body. If the point on this sheet of paper that has zero velocity in the lab frame is outside of the original rigid body, it is an ICoR that is fixed to the frame of the rigid body but external to it. Hence, Eq. (3.25) is still valid when using this point as a reference point as it is on the "extension" of the original rigid body.

To determine the ICoR of a rigid body, either the velocities of two points on the rigid body or the combination of the velocity of one point and the angular velocity of the body is required. The crux in locating the ICoR lies upon the fact that the vector joining it to a point on a rigid body must be perpendicular to the instantaneous velocity of that point (as $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$). In all cases, the angular velocity of the body can be computed as well — enabling us to determine the velocities of all points on the rigid body. Let us consider the following two-dimensional cases:

Parallel Velocities of Two Points on the Body

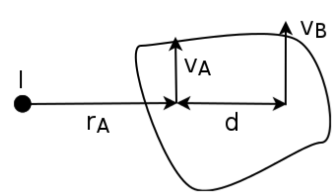


Figure 3.17: Parallel velocities

In such situations, the velocities of two points on the body are given and are known to be parallel. They are also perpendicular to the line joining them. Then, the ICoR must be along the line joining the two velocities as the position vector of these two points relative to the ICoR is perpendicular to their velocities, which are along the same direction. Without loss of generality, let the velocity of B, v_B be larger than that of A, v_A. Let r_A be the distance between the ICoR and point A and let the distance between A and B be d. Note that the ICoR lies on the side closer to A. As $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$ and $\boldsymbol{\omega}$ is the same for both points,

$$\frac{v_A}{r_A} = \frac{v_B}{r_A + d}$$

$$r_A = \frac{v_A d}{v_B - v_A}$$

$$\omega = \frac{v_B - v_A}{d}.$$

Non-parallel Velocities of Two Points on the Body

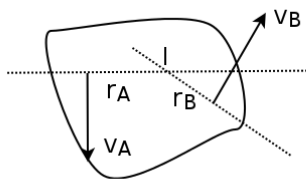


Figure 3.18: Non-parallel velocities

When two non-parallel velocities of two points on a rigid body are given, the ICoR is the point of intersection of the two lines perpendicular to the velocity of the two points — it is the only point from which emanating vectors to the two points are perpendicular to the corresponding velocities. Then, the two distances between the two points and the ICoR, r_A and r_B can be computed via geometric means. Consequently, the angular velocity is

$$\omega = \frac{v_A}{r_A} = \frac{v_B}{r_B}.$$

The direction of $\boldsymbol{\omega}$ (clockwise or anti-clockwise) has to be determined by considering the direction of the velocity of either point and that of the vector joining the ICoR to that particular point.

Velocity of One Point on the Body and the Angular Velocity

When the velocity of a certain point A and the angular velocity of the entire rigid body are given, the ICoR again lies along the line perpendicular to A's velocity. Then, we can find the position of the ICoR from the fact that

$$\boldsymbol{v}_A = \boldsymbol{\omega} \times \boldsymbol{r}_A.$$

The distance between A and the ICoR can be easily computed as

$$r_A = \frac{v}{\omega_A}.$$

There are two points along the perpendicular that satisfy this. The correct location is determined by considering the cross product.

For a three-dimensional object, define the xy-plane to be perpendicular to $\boldsymbol{\omega}$. Then, if we are given any of the above properties of two points, consider their projections on the xy-plane such that they are essentially on the same plane. Then, the above analysis can be applied once again as one can easily show that the z-coordinate of those points are inconsequential when applying $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$. Once we have found the ICoR, all points along the axis parallel to $\boldsymbol{\omega}$ and passing through the ICoR in fact have zero velocity in the lab frame. Thus, we have an instantaneous axis of rotation in the three-dimensional case.

Lastly, note that in most cases, the ICoR changes with the motion of the rigid body. This is true even in the case when the velocity of the reference point and the angular velocity of the rigid body remains constant. This is due to the fact that the object is physically translating through space. Hence, the utility of the ICoR only lies in determining the velocity of a general point on the body at a particular instant.

Problem: Solve the previous problem again by considering the ICoR.

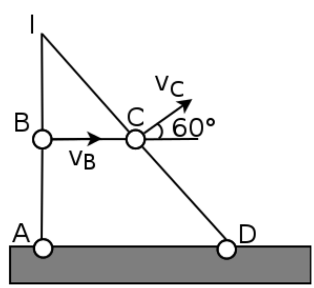


Figure 3.19: ICoR of rod BC

Since the velocities of B and C are constrained to be perpendicular to rods AB and CD respectively, the ICoR, which lies on the intersection of the two lines perpendicular to the velocities of two points on a rigid body, is the point of intersection I obtained from extending lines AB and DC. Now, the angular velocity vector of rod BC, $\boldsymbol{\omega}_{BC}$ can be computed as

$$\boldsymbol{\omega}_{BC} = \frac{v_B}{\overline{IB}}\hat{\boldsymbol{k}} = \frac{\omega l}{\overline{IB}}\hat{\boldsymbol{k}}$$

pointing out of the page. However, note that $\boldsymbol{\omega}$ of rod AB is pointing into the page. Use the right-hand grip rule to verify these if necessary. The velocity of point C, v_C, is perpendicular to line IC and is of magnitude

$$v_C = \overline{IC} \cdot \omega_{BC} = \frac{\overline{IC}}{\overline{IB}} \cdot \omega l = \frac{\omega l}{\sin 30°} = 2\omega l.$$

The vector $\boldsymbol{v}_C$ makes an angle of 60° with the positive x-axis. Hence,

$$\boldsymbol{v}_C = l\omega\hat{\boldsymbol{\imath}} + \sqrt{3}l\omega\hat{\boldsymbol{\jmath}}.$$

Additive Property of Angular Velocities

It is beneficial to underscore the additive property of angular velocities here, once and for all. Let a rigid body B have angular velocities $\boldsymbol{\omega}_{B1}$ and $\boldsymbol{\omega}_{B2}$ relative to frames 1 and 2 which have origins O and O' respectively (that do not necessarily coincide). If the axes of frame 2 have angular velocity $\boldsymbol{\omega}_{21}$ in frame 1, we claim that

$$\boldsymbol{\omega}_{B1} = \boldsymbol{\omega}_{B2} + \boldsymbol{\omega}_{21}. \tag{3.26}$$

Proof: A major part of our proof would rely on the following theorem. If a point P has velocities $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ as observed in frames 1 and 2, with the axes of frame 2 rotating at angular velocity $\boldsymbol{\omega}_{21}$ relative to frame 1, and if the instantaneous position vector of P relative to the origin of frame 1 is $\boldsymbol{r}$, we have

$$\boldsymbol{v}_1 = \boldsymbol{v}_2 + \boldsymbol{\omega}_{21} \times \boldsymbol{r}. \tag{3.27}$$

The origins of frames 1 and 2 need not coincide. This equation has a haunting resemblance to Eq. (3.25) but the reader should understand that these are two vastly different equations. Equation (3.25) describes the velocity of a point on a rigid body and, at most, describes the transformation of the velocity of that point as seen by a reference particle to that in the lab frame (the axes of these frames do not rotate relative to each other) while Eq. (3.27) represents the transformation of velocities across relatively rotating frames. This transformation is evidently necessary in this case as axes that are fixed to frame 2 are now rotating with respect to frame 1. Actually, the tools required to prove this equation have already been developed but readers who wish to directly read the proof can skip ahead to Section 11.2.

Proceeding with the actual proof, pick an arbitrary point P on the rigid body and let the instantaneous center of rotation[11] of the rigid body in frame 1 be R. Choosing R as the reference point to apply Eq. (3.25) to, the velocity

[11] Actually R can be any point on the rigid body and this proof would still hold. However, choosing it as the ICoR simplifies matters here.

of P in frame 1 is

$$\boldsymbol{v}_{P1} = \boldsymbol{\omega}_{B1} \times \boldsymbol{r}_{RP},$$

where $\boldsymbol{r}_{RP}$ is the vector pointing from R to P. Meanwhile, the velocity of P in frame 2 is

$$\boldsymbol{v}_{P2} = \boldsymbol{v}_{R2} + \boldsymbol{\omega}_{B2} \times \boldsymbol{r}_{RP}$$

by Eq. (3.25), where $\boldsymbol{v}_{R2}$ is the velocity of R as observed in frame 2. $\boldsymbol{v}_{R2}$ can be computed by applying Eq. (3.27).

$$\boldsymbol{v}_{R1} = \boldsymbol{v}_{R2} + \boldsymbol{\omega}_{21} \times \boldsymbol{r}_{OR},$$

where $\boldsymbol{v}_{R1} = \boldsymbol{0}$ is the velocity of R as observed in frame 1 and $\boldsymbol{r}_{OR}$ is the position vector of R in frame 1. Therefore,

$$\boldsymbol{v}_{P2} = \boldsymbol{\omega}_{B2} \times \boldsymbol{r}_{RP} - \boldsymbol{\omega}_{21} \times \boldsymbol{r}_{OR}.$$

Finally, we can relate $\boldsymbol{v}_{P1}$ and $\boldsymbol{v}_{P2}$ by applying Eq. (3.27) again.

$$\boldsymbol{v}_{P1} = \boldsymbol{v}_{P2} + \boldsymbol{\omega}_{21} \times \boldsymbol{r}_{OP},$$

where $\boldsymbol{r}_{OP}$ is the position vector of P in frame 1. Combining these expressions, we obtain

$$\boldsymbol{\omega}_{B1} \times \boldsymbol{r}_{RP} = \boldsymbol{\omega}_{B2} \times \boldsymbol{r}_{RP} + \boldsymbol{\omega}_{21} \times (\boldsymbol{r}_{OP} - \boldsymbol{r}_{OR})$$
$$\implies \boldsymbol{\omega}_{B1} \times \boldsymbol{r}_{RP} = (\boldsymbol{\omega}_{B2} + \boldsymbol{\omega}_{21}) \times \boldsymbol{r}_{RP}.$$

Since P is arbitrary and $\boldsymbol{r}_{RP}$ follows suit, we must have

$$\boldsymbol{\omega}_{B1} = \boldsymbol{\omega}_{B2} + \boldsymbol{\omega}_{21}.$$

Problem: A toy consists of some blades attached to a common center O'. The blades rotate at an angular speed ω_2 about the center. Now, the center of the toy is constrained to undergo uniform circular motion about a center O at anti-clockwise angular velocity $\boldsymbol{\omega}_1$, while its blades rotate about the center O' at anti-clockwise angular velocity $\boldsymbol{\omega}_2$. If the position vector of O' with respect to O is $\boldsymbol{R}$ at this instance, determine the velocity of a point on the blades that has position vector $\boldsymbol{r}$ relative to the center O' in the lab frame in which O is stationary.

The important observation here is that the angular velocity of the blades in the lab frame is $\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2$ and not $\boldsymbol{\omega}_2$. To convince yourself that this is the case, consider the specific scenario where $\boldsymbol{\omega}_2 = \boldsymbol{0}$ and the center O' simply rotates at $\boldsymbol{\omega}_1$ about O. Mark a point P on the blade — you will observe that when O' has rotated an angle θ, P would have also rotated by θ (refer to the question pertaining to Fig. 3.23 for an illustration).

In this scenario, frame 2 is a set of coordinate axes fixed to O' that will rotate with O' at angular velocity $\boldsymbol{\omega}_1$ relative to the lab frame (we define this as frame 1). Since the angular velocity of the blades as observed in frame 2 is $\boldsymbol{\omega}_2$, the angular velocity $\boldsymbol{\omega}$ of the blades with respect to frame 1 is then

$$\boldsymbol{\omega} = \boldsymbol{\omega}_1 + \boldsymbol{\omega}_2.$$

Applying Eq. (3.25) and choosing the center O', which travels instantaneously at velocity $\boldsymbol{\omega}_1 \times \boldsymbol{R}$, as the reference point yields the velocity of a point on the blade at a position $\boldsymbol{r}$ from O' as

$$\boldsymbol{v} = \boldsymbol{\omega}_1 \times \boldsymbol{R} + (\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \times \boldsymbol{r},$$

in the lab frame. For those who are still confused about why we could not directly add the velocity of the point as seen in frame 2, $\boldsymbol{\omega}_2 \times \boldsymbol{r}$, to the velocity of the origin O' of frame 2 as observed in frame 1, $\boldsymbol{\omega}_1 \times \boldsymbol{R}$, stand up and spin on the spot. Objects in your room will appear to possess a certain velocity in your frame (frame 2) but you know that a friend (frame 1) who is not rotating would observe them to be stationary, even though your friend observes that you remain on the same spot on the ground — contradicting the previous statement! The resolution here is that we have applied the additive property of velocities wrongly. It is right to say that the velocity of a point in frame 1 is the addition of the velocity of O' in frame 1 and the velocity of that point relative to O' as observed in **frame 1**[12] but it is fallacious to say that it is the addition of the velocity of O' in frame 1 and the velocity of that point relative to O' as observed in **frame 2**, which is rotating relative to frame 1. The moral of the story is that a real danger lurks in rotating frames and we shall explore how to tackle them in Section 11.2.

Angular Acceleration

The angular acceleration vector of a rigid body is defined to be the rate of change of the angular velocity vector.

$$\boldsymbol{\alpha} = \frac{d\boldsymbol{\omega}}{dt}. \tag{3.28}$$

Since $\boldsymbol{\omega}$ is a property of the entire body that is invariant across non-rotating frames, $\boldsymbol{\alpha}$ follows suit. Like any other time derivative of a vector, $\boldsymbol{\alpha}$ can

[12] Any frame that is non-rotating relative to frame 1 works too.

change in two ways — a change in the magnitude ω or the direction $\hat{\boldsymbol{\omega}}$. Rewriting $\boldsymbol{\alpha}$ in a more suggestive way,

$$\boldsymbol{\alpha} = \frac{d\omega\hat{\boldsymbol{\omega}}}{dt} = \dot{\omega}\hat{\boldsymbol{\omega}} + \omega\frac{d\hat{\boldsymbol{\omega}}}{dt}.$$

In this book, only fixed axis rotations will be analyzed. That is, the direction of the angular velocity vector does not change. Then, the second term is zero and

$$\boldsymbol{\alpha} = \dot{\omega}\hat{\boldsymbol{\omega}} = \alpha\hat{\boldsymbol{\omega}}.$$

Finally, the acceleration of an arbitrary point on a rigid body, relative to that of a reference point can be obtained by differentiating Eq. (3.25),

$$\frac{d\boldsymbol{v}}{dt} = \frac{d\boldsymbol{v}_{ref}}{dt} + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r} + \boldsymbol{\omega} \times \frac{d\boldsymbol{r}}{dt}.$$

Note that $\frac{d\boldsymbol{r}}{dt}$ is the rate of change of the vector joining the reference point to the point of concern, which is $\boldsymbol{v} - \boldsymbol{v}_{ref} = \boldsymbol{\omega} \times \boldsymbol{r}$. Hence, the instantaneous acceleration of an arbitrary point on the body, $\boldsymbol{a}$, in relation to the instantaneous acceleration of the reference point, $\boldsymbol{a}_{ref}$ is

$$\boldsymbol{a} = \boldsymbol{a}_{ref} + \boldsymbol{\alpha} \times \boldsymbol{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}), \tag{3.29}$$

where $\boldsymbol{\omega}$ and $\boldsymbol{\alpha}$ are the instantaneous angular velocity and acceleration vectors, respectively. This equation is entirely general and is not restricted to fixed axis rotations. In the case of fixed axis rotations,

$$\boldsymbol{a} = \boldsymbol{a}_{ref} + \alpha\hat{\boldsymbol{\omega}} \times \boldsymbol{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}).$$

If the translational motion and the plane of rotation of the rigid body strictly lie in a two-dimensional plane, one can evaluate the cross products easily as $\boldsymbol{r}$ and $\boldsymbol{\omega}$ are perpendicular, where $\boldsymbol{r}$ points from the reference point to the point of concern. It is easy to verify that $\hat{\boldsymbol{\omega}} \times \boldsymbol{r} = r\hat{\boldsymbol{\theta}}$ while the last term can be simplified via the BAC-CAB rule (Eq. (3.6)).

$$\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) = \boldsymbol{\omega}(\boldsymbol{\omega} \cdot \boldsymbol{r}) - \boldsymbol{r}(\boldsymbol{\omega} \cdot \boldsymbol{\omega}) = -\omega^2\boldsymbol{r},$$

since $\boldsymbol{\omega} \cdot \boldsymbol{r} = \boldsymbol{0}$. Applying these simplifications yields

$$\boldsymbol{a} - \boldsymbol{a}_{ref} = r\alpha\hat{\boldsymbol{\theta}} - r\omega^2\hat{\boldsymbol{r}}.$$

Similarly, splitting the terms into the radial and tangential directions,

$$a_r - a_{ref,r} = -r\omega^2,$$

$$a_t - a_{ref,t} = r\alpha.$$

That is, the relative radial acceleration must provide the necessary centripetal acceleration for the particle of concern to continue traveling at a constant distance away from the reference point. Furthermore, the relative tangential acceleration must reflect the angular acceleration.

Problem: With regards to the previous problem, if rod AB has an angular acceleration vector $\boldsymbol{\alpha}_{AB} = -\alpha\hat{\boldsymbol{k}}$, determine the acceleration of point C. Rod A still has an angular velocity vector $\boldsymbol{\omega}_{AB} = -\omega\hat{\boldsymbol{k}}$. The lengths of AB, BC and CD are l, $2l$ and $2l$ respectively.

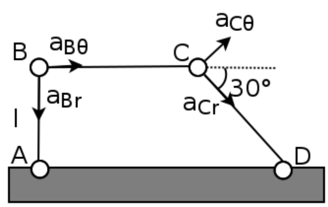


Figure 3.20: Acceleration of points

Consider the components of the accelerations of points B and C as labeled in Fig. 3.20 above. The respective accelerations along rods AB and CD are simply the centripetal acceleration terms.

$$a_{Br} = \frac{v_B^2}{l} = \omega^2 l,$$

$$a_{Cr} = \frac{v_C^2}{l} = 4\omega^2 l.$$

From the kinematics equation in polar coordinates,

$$a_{B\theta} = l\alpha.$$

Lastly, we need to determine $a_{C\theta}$ by imposing the condition that the relative acceleration of points B and C along rod BC must provide the necessary centripetal acceleration. Thus, the relative velocity between points B and C in the direction perpendicular to rod BC, or the angular velocity of rod BC, must first be determined. The former shall be utilized as it is simply v_{Cy} that we have calculated previously. As the relative acceleration between B and C along rod BC must correspond to the necessary centripetal acceleration,

$$\begin{aligned} a_{B\theta} - a_{C\theta}\cos 60^\circ &= \frac{v_{Cy}^2}{2l} = \frac{3}{2}l\omega^2 \\ a_{C\theta} &= 2l\alpha - 3l\omega^2. \end{aligned}$$

Having determined a_{Cr} and $a_{c\theta}$, the acceleration of point C can be resolved into its x and y components.

$$\boldsymbol{a}_C = \left(l\alpha + \frac{4\sqrt{3}-3}{2}l\omega^2\right)\hat{\boldsymbol{i}} + \left(\sqrt{3}l\alpha - \frac{4+3\sqrt{3}}{2}l\omega^2\right)\hat{\boldsymbol{j}}.$$

3.5.2 *Rolling*

Consider a rigid circular wheel of radius R exhibiting both translational motion in the x-direction and rotational motion on a surface. It is rotating clockwise at an angular velocity ω (i.e. the angular velocity vector is $-\omega\hat{\boldsymbol{k}}$ with the z-axis pointing out of the page).

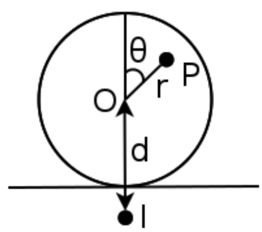


Figure 3.21: Rolling wheel

Usually, the center of the wheel is a pivotal reference point in describing its motion. Let the velocity of the center be v. Then, the instantaneous center of rotation is a distance

$$d = \frac{v}{\omega},$$

below the center of the wheel. Then, the velocity of an arbitrary point on the wheel can be determined. Consider a point P on the wheel, a distance r, $r \le R$, away from the center of the wheel, O and at an angular coordinate θ (Fig. 3.21). Defining the x and y-axes to be positive rightwards and upwards, the position vector of P relative to the instantaneous center of rotation, I, is given by

$$\boldsymbol{r'} = \begin{pmatrix} r\sin\theta \\ \frac{v}{\omega} + r\cos\theta \\ 0 \end{pmatrix}.$$

The angular velocity vector of the body is $\begin{pmatrix} 0 \\ 0 \\ -\omega \end{pmatrix}$ — the negative sign arises from the fact that ω is defined to be clockwise. Then, the velocity of point P is

$$\boldsymbol{v}' = \begin{pmatrix} 0 \\ 0 \\ -\omega \end{pmatrix} \times \begin{pmatrix} r\sin\theta \\ \frac{v}{\omega} + r\cos\theta \\ 0 \end{pmatrix} = \begin{pmatrix} v + r\omega\cos\theta \\ -\omega r\sin\theta \\ 0 \end{pmatrix}. \tag{3.30}$$

This formula can also be obtained from applying Eq. (3.25) which effectively apportions the net velocity of point P to contributions by the velocity of the center of the circle and the component due to the rotation of P about O. This will be illustrated diagrammatically in the following scenario.

Rolling Without Slipping

Slipping occurs when there is relative motion between surfaces. Hence, there must be no relative velocity and acceleration between the corresponding points of contact on the object and on the surface for there to be no slipping. Naturally, failure in satisfying such conditions engenders slipping.

Consider the set-up in the previous section with a stationary ground. Rather than applying Eq. (3.25) directly, we shall decompose the motion of a point on the body into a translation of O and a rotation of the point about O. Let us just consider the top and bottom of the circle for the sake of illustration.

All points on the wheel move at a velocity v towards the right due to the translation of the wheel with respect to the center. This, in addition to the rotational component of the motion of the point, determined by the distance between the center and the point and the point's angular position relative to the object, fully describes the motion of each individual point. We notice that the top of the wheel travels at velocity $v + R\omega$ while the bottom travels at velocity $v - R\omega$ (Fig. 3.22).

Slipping occurs when two surfaces have a non-zero relative velocity with respect to one another. Thus, assuming that the surface is stationary at the bottom of the wheel, the wheel rolls without slipping when the bottom

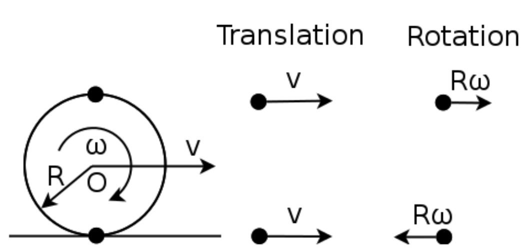


Figure 3.22: Rolling without slipping

of wheel has zero velocity (i.e. $v = R\omega$) and slips whenever $v > R\omega$ or $v < R\omega$. Note that we do not need to assume that the surface is flat (as depicted in Fig. 3.22) as we just require the circle to roll without slipping instantaneously. Hence, the non-slip condition is

$$v = R\omega.$$

To obtain the relationship between the acceleration of O, a, and the angular acceleration, defined to be α clockwise, the above equation can be differentiated with respect to time. That is,

$$a = R\alpha.$$

Lastly, performing an integration on the previous equation with respect to time, the distance travelled by the center of the circle is related to the angular distance according to

$$d = R\Delta\theta,$$

if the circle rolls without slipping throughout the motion.

Problem: A coin of radius r rolls without slipping on the interior of a hollow circle of radius kr for one revolution and returns to its original position. How many rounds did the coin rotate about its center?

Well, one might think that since the contact point between the coin and the circle travels a distance $2\pi kr$ and because the circumference of the coin is $2\pi r$, the coin must have rotated k rounds.

However, notice that the center of the coin only traveled a distance $2\pi(k-1)r$ (i.e. in a circle of radius $(k-1)r$). Hence the coin only rotates $k-1$ rounds around its center.

There is another way to understand this. Without the loss of generality, assume that the coin revolves around the circle in the clockwise direction. Then it must rotate anti-clockwise to prevent slipping. If the coin traveled on a flat surface for a distance $2\pi kr$, it rotates k rounds anti-clockwise. However, in this case, even if the coin simply translated with a fixed contact point (represented by the black dot in Fig. 3.23) clockwise in the interior of the hollow circle, it would have rotated one clockwise round about its center. Since, the motion of the coin is the composition of both translational and rotational components, the coin rotates a total of $k-1$ rounds anti-clockwise.

Problem: A spool consists of a cylindrical axle of radius r and a larger cylindrical rim of radius R. A weightless rope is wound around the axle. The loose end of the rope is pulled at this instant and has a velocity v_0 in the

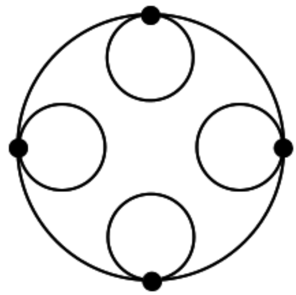


Figure 3.23: Additional clockwise round

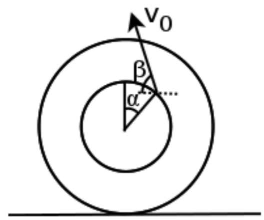


Figure 3.24: Pulling a spool

configuration shown in Fig. 3.24. If the spool rolls without slipping on the flat ground, what is the velocity of the center of the spool at this instant?

The velocity of the point of intersection between the loose part of the rope and the axle is obtained by substituting α for θ in Eq. (3.30),

$$\boldsymbol{v'} = \begin{pmatrix} v + r\omega\cos\alpha \\ -r\omega\sin\alpha \\ 0 \end{pmatrix}.$$

Then, the component of this velocity along the direction of the rope is

$$-v'_x\cos\beta + v'_y\sin\beta = -(v + r\omega\cos\alpha)\cos\beta - r\omega\sin\alpha\sin\beta = v_0.$$

Lastly, the non-slip condition implies that $v = R\omega$. Note that the rim is a distance R from the center and not r. Then

$$-v\cos\beta - \frac{r}{R}v\cos(\alpha-\beta) = v_0$$

$$v = \frac{-v_0}{\frac{r}{R}\cos(\alpha-\beta)+\cos\beta}.$$

The negative sign indicates that the center of the spool travels in the negative x-direction (towards the left).

3.5.3 *Constrained Motion*

In some problems, the motion of a rigid body is constrained to move in a certain manner. These restrictions provide additional equations that must be satisfied by the system and often reduce the number of coordinates needed to define a unique state of the system. Consider the following ubiquitous set-up.

Kinematics of a Leaning Ladder

Consider the motion of a rigid ladder of length $2l$ that is leaning on an impenetrable vertical wall with its leg supported by a horizontal floor.

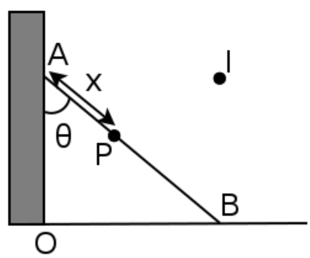


Figure 3.25: Leaning ladder

Evidently, the ladder is constrained to move along the surfaces of the wall and the floor. Hence, there must be a relationship between its angular velocity and the velocity of each point on the ladder. This relationship can be easily determined by finding the ICoR of the ladder. As the velocity of the ends of the ladder are restricted to be along the surface of the floor and ground, the ICoR, denoted by I, is the point of intersection of the lines perpendicular to these surfaces. Now let $\boldsymbol{\omega}$ be the angular velocity of the ladder. Then,

$$\boldsymbol{\omega} = \dot{\theta}\hat{\boldsymbol{k}}.$$

The positive sign in the equation above arises from the fact that an anti-clockwise rotation of the ladder increases θ (recall that an angular velocity vector in the positive z-direction represents an instantaneous anti-clockwise rotation). If we had defined θ to be $\angle ABO$ instead, a negative sign must be included in front of $\dot{\theta}$. Now, consider a particular point P on the ladder that is at a distance x from the top of the ladder. Its coordinates are $(x\sin\theta,$ $(2l-x)\cos\theta)$. The coordinates of I are $(2l\sin\theta, 2l\cos\theta)$. Therefore,

$$\vec{\boldsymbol{IP}} = \begin{pmatrix} x\sin\theta \\ (2l-x)\cos\theta \\ 0 \end{pmatrix} - \begin{pmatrix} 2l\sin\theta \\ 2l\cos\theta \\ 0 \end{pmatrix} = \begin{pmatrix} (x-2l)\sin\theta \\ -x\cos\theta \\ 0 \end{pmatrix}.$$

Hence, the velocity of P is

$$\boldsymbol{v} = \begin{pmatrix} 0 \\ 0 \\ \dot{\theta} \end{pmatrix} \times \begin{pmatrix} (x-2l)\sin\theta \\ -x\cos\theta \\ 0 \end{pmatrix} = \begin{pmatrix} x\cos\theta\dot{\theta} \\ (x-2l)\sin\theta\dot{\theta} \\ 0 \end{pmatrix}.$$

As we will be dealing with the center of a ladder frequently (especially if the ladder has a uniform mass distribution), let us compute its velocity in terms of $\dot{\theta}$. Substituting $x = l$,

$$\boldsymbol{v}_{center} = \begin{pmatrix} l\cos\theta\dot{\theta} \\ -l\sin\theta\dot{\theta} \\ 0 \end{pmatrix}.$$

This expression can also be obtained by dividing the velocity of the ends of the ladder into translational and rotational components while using the center of the ladder as the reference point.

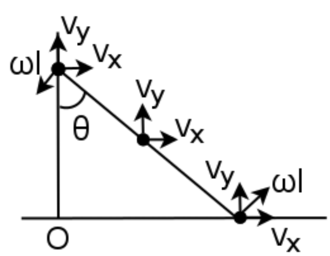


Figure 3.26: Components of velocities

Let the center of the ladder possess horizontal and vertical components of velocity v_x and v_y. The velocity of the ends of the ladder is the composition of the velocity of the center of the ladder and their corresponding velocities due to their rotation about the center (Fig. 3.26). In order for the velocity of the top end of the ladder to be strictly along the wall,

$$v_x = l\omega\cos\theta = l\cos\theta\dot{\theta}.$$

Similarly, for the bottom end of the ladder to remain on the ground,

$$v_y = -l\omega\sin\theta = -l\sin\theta\dot{\theta}.$$

It is then natural to ponder about the dependence of the acceleration of the center of the ladder on $\ddot{\theta}$. However, it is much more convenient to differentiate the above equations with respect to time than to consider the accelerations of the ends of the ladder, via Eq. (3.29) relative to the center of the ladder. This will be elaborated on in the following section.

Obtaining Relationships between Kinematic Quantities via Differentiation

Often, when there are constraints on a system, it is much easier to obtain the relationship between certain kinematic quantities and their time derivatives at all instances by progressively differentiating the equation governing the restriction imposed by the constraints. Referring to the previous situation pertaining to the leaning ladder, the coordinates of the center of the ladder are

$$\boldsymbol{r}_{center} = \begin{pmatrix} l\sin\theta \\ l\cos\theta \end{pmatrix}.$$

Taking the derivative of this with respect to time,

$$\boldsymbol{v}_{center} = \begin{pmatrix} l\cos\theta\dot{\theta} \\ -l\sin\theta\dot{\theta} \end{pmatrix},$$

which is consistent with our result in the previous section. Further differentiating this with respect to time, the acceleration of the center of the ladder as a function of θ and its time derivatives is

$$\boldsymbol{a}_{center} = \begin{pmatrix} -l\sin\theta\dot{\theta}^2 + l\cos\theta\ddot{\theta} \\ -l\cos\theta\dot{\theta}^2 - l\sin\theta\ddot{\theta} \end{pmatrix}.$$

Note that in the steps above, we have implicitly used the fact that the basis vectors, $\hat{\boldsymbol{\imath}}$ and $\hat{\boldsymbol{\jmath}}$, are immutable. If this were not the case, the time derivatives of these basis vectors must also be included, in a fashion similar to the derivation of the kinematics equations in polar coordinates. Lastly, there is also an interesting geometrical property of this set-up — the motion of the center of the ladder traces out a circular arc of radius l about the origin O. This is evident from $\boldsymbol{r}_{center}$ which implies that the center of the ladder is at a constant distance $\sqrt{l^2\sin^2\theta + l^2\cos^2\theta} = l$ from the origin.

Next, let us consider a scenario in which objects are to remain in contact. Three identical cylinders of radius R are initially held in contact on the ground in the configuration shown below.

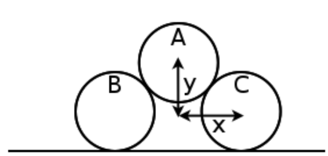


Figure 3.27: Three cylinders

The cylinders are then released. If cylinder A remains in contact with B and C for a certain period of time, what is the relationship between the velocities of the cylinders and their displacements during this period?

The resultant velocity of cylinder A is solely along the vertical direction due to symmetry. Define a coordinate system as depicted in the figure above. Let y be the y-coordinate of cylinder A and let x be the x-coordinate of cylinder C. Then, for cylinders A and C to remain in contact,

$$x^2 + y^2 = 4R^2.$$

Differentiating this equation with respect to time,

$$2x\dot{x} + 2y\dot{y} = 0$$

$$x\dot{x} = -y\dot{y}.$$

This is the equation relating the velocities of the cylinders and their displacements. Note that the displacement and velocity of cylinder B are simply $-x$ and $-\dot{x}$ respectively due to symmetry. Next, we can differentiate the above equation with respect to time again to extract information about further derivatives. Here,

$$\dot{x}^2 + x\ddot{x} = -\dot{y}^2 - y\ddot{y}.$$

All-in-all, this method of differentiating equations to relate certain variables is extremely useful in determining their relationships at all instances in time. However, in the case where the relationship between certain quantities are needed at only a particular instant, akin to the example of the connected rigid rods, this method may prove to be somewhat tedious, especially for systems with a large number of degrees of freedom. In such cases, it is expeditious to use the kinematics equations such as Eqs. (3.25) and (3.29). Ultimately, the equations in this chapter are not physically insightful by themselves, as they are mostly mathematical consequences that have little relation to the laws of physics. The dynamical laws, which stem from observations and deductions of the world, provide the equation of motion, after which the problem is merely a kinematic one. However, the study of kinematics and the inherent relationships between quantities in a system with constraints provides valuable equations which are paramount in solving for the system of equations obtained from the dynamical laws.

3.6 Univariate Differential Equations

This section will cover the tricks in solving various differential equations involving a single variable that is a function of an independent parameter — an essential component in solving for the kinematic quantities, after the dynamical laws have been applied. For the rest of this section, we hope to solve for a dependent variable $y(x)$ in terms of an independent variable x. $\frac{dy}{dx}$ will be denoted as y'. Though many physical examples will be included, note that this section will be purely mathematical and is hence highly general.

3.6.1 *Separable Differential Equations*

A separable differential equation takes the form

$$y' = f(x),$$

where $f(x)$ is some arbitrary function of x. As implied by the term "separable", we can solve for y by shifting the dx term in the denominator of y' to the right-hand side.

$$dy = f(x)dx.$$

Integrating, we solve for y.

$$y = \int f(x)dx + c,$$

where the constant of integration c, also captures the constant of integration from $\int f(x)dx$, for the sake of convenience.

One-Dimensional Motion with Linear Drag

Though we have not formally introduced Newton's laws, consider the one-dimensional motion of an object under the influence of a linear drag force and a constant external force. Its initial velocity is in the same direction as the external force. Then, its equation of motion will be of the form

$$m\boldsymbol{a} = m\boldsymbol{A} - bv\hat{\boldsymbol{v}},$$

where $\boldsymbol{A}$ is a constant vector which represents the constant acceleration that the object will undergo if there was no drag. Aligning the x-axis along the line of motion, we can rewrite the vector equation above in terms of a single scalar variable, the x-coordinate of the object. Let a and v denote the acceleration and velocity of the particle, respectively. Then,

$$a = A - kv,$$

where $k = \frac{b}{m}$. Using the fact that $a = \frac{dv}{dt}$,

$$\frac{dv}{dt} = A - kv.$$

Observe that this differential equation is separable. Separating the variables and integrating,

$$\int_{v_0}^{v} \frac{1}{v - \frac{A}{k}} dv = \int_0^t -k dt$$

$$\ln \left| \frac{v - \frac{A}{k}}{v_0 - \frac{A}{k}} \right| = -kt,$$

where v_0 is the initial velocity. To remove the absolute value signs, observe that if the initial velocity $v_0 \geq \frac{A}{k}$, then $v \geq \frac{A}{k}$ at all later instances as a, which is the signed magnitude of acceleration, is negative until v attains the value $\frac{A}{k}$, after which the acceleration will be zero. A similar logic can be used to deduce that if $v_0 \leq \frac{A}{k}$, then $v \leq \frac{A}{k}$ at all later instances. Hence, we can simply remove the absolute value sign as the numerator and the denominators are ensured to be of the same sign. Then,

$$v = \left(v_0 - \frac{A}{k}\right) e^{-kt} + \frac{A}{k}.$$

Notice that as time tends to infinity, the object reaches the velocity $\frac{A}{k}$, which is known as the terminal velocity of the object. Next, the equation above can be integrated again to solve for the x-coordinate of the particle as a function of time.

$$x = \left(\frac{A}{k^2} - \frac{v_0}{k}\right)\left(e^{-kt} - 1\right) + \frac{A}{k}t + x_0,$$

where x_0 is the initial x-coordinate of the object. An interesting result occurs when $A = 0$ (i.e. there is no other external force besides the drag force) as the total distance traveled by the object converges to $\frac{v_0}{k}$, though the process takes an indefinitely long time.

One-Dimensional Motion with Quadratic Drag

Consider an object that travels solely along the x-direction, under the sole influence of a quadratic drag force. Its equation of motion is

$$a = -kv^2,$$

where $k = \frac{c}{m}$ and c is the constant in $\boldsymbol{F}_{quad}$. Writing a as $\frac{dv}{dt}$ and separating variables,

$$\int_u^v \frac{1}{v^2} dv = \int_0^t -k dt$$

$$\frac{1}{u} - \frac{1}{v} = -kt$$

$$v = \frac{u}{1 + ukt}.$$

Separating and integrating again,

$$\int_{x_0}^{x} dx = \int_0^t \frac{u}{1 + ukt} dt$$

$$x - x_0 = u \ln|1 + ukt|$$

$$x = u \ln|1 + ukt| + x_0.$$

It is rather interesting to compare this result with that of the linear case without a constant acceleration. As $t \to \infty$, both velocities tend to zero. However, the velocity in the linear case decays much faster (exponentially) as compared to that in the quadratic case which is approximately inversely proportional[13] to t for large t.

This subtle difference manifests itself in the x-coordinate of the object as a function of time. As $t \to \infty$, the linear case produces a convergent result. However, in this quadratic case, the x-coordinate of the particle also tends to infinity, implying that it is not "bounded". This discrepancy can be understood qualitatively. As the speed of the object becomes small, v^2 is even smaller, causing the deceleration of the object to be smaller in magnitude than the linear case.

3.6.2 *Making Equations Separable*

Since separable equations can be solved directly, the goal of this section is to express differential equations in a more illuminating form via certain tricks.

y'' in terms of y

Consider a differential equation of the form

$$y'' = f(y).$$

Observe that

$$y'' = \frac{dy'}{dx} = \frac{dy'}{dy} \cdot \frac{dy}{dx} = y' \frac{dy'}{dy}. \tag{3.31}$$

[13] As t is large, we can simply ignore the additional term of 1 in the denominator of v.

Then,

$$y' \frac{dy'}{dy} = f(y).$$

This is a separable equation! Separating variables and integrating,

$$\int y' dy' = \int f(y) dy$$
$$y'^2 = 2\int f(y) dy + c,$$

where c is a constant of integration. If we are given the exact function $f(y)$, we can square root both sides of the above equation and separate variables again to solve for $x(y)$ and therefore, $y(x)$, if $x(y)$ can be inverted. Finally, note that another way of expressing y'' is $\frac{1}{2}\frac{dy'^2}{dy}$ as $y'dy' = \frac{1}{2}d(y'^2)$.

Integrating Factor

A rather ubiquitous differential equation takes the form

$$y' + f(x)y = g(x). \tag{3.32}$$

Here's the idea behind solving this differential equation. Suppose that we could find a function $h(x)$ such that when both sides of the equation are multiplied by $h(x)$, the left-hand side becomes

$$\frac{d(h(x)y)}{dx}.$$

The entire equation will become

$$\frac{d(h(x)y)}{dx} = g(x)h(x),$$

which is a separable differential equation. Separating the variables and integrating,

$$\int d(h(x)y) = \int g(x)h(x)dx$$
$$h(x)y = \int_{x_0}^{x} g(x)h(x)dx + h(x_0)y_0$$
$$y = \frac{\int_{x_0}^{x} g(x)h(x)dx + h(x_0)y_0}{h(x)}. \tag{3.33}$$

Hence, the only task left is to find such a function $h(x)$, which is known as the integrating factor. Observe that

$$\frac{d(h(x)y)}{dx} = h(x)y' + h'(x)y.$$

Comparing this to the left-hand side of the original equation multiplied by $h(x)$,

$$h(x)y' + h'(x)y = h(x)y' + h(x)f(x)y.$$

Hence,

$$\frac{h'(x)}{h(x)} = f(x).$$

Making the astute observation that the left-hand side of this equation is $\frac{d(\ln h(x))}{dx}$,

$$\int d(\ln h(x)) = \int f(x)dx$$
$$\ln h(x) = \int f(x)dx + c,$$

where we have also "pulled out" the constant of integration from the $\int f(x)dx$ and absorbed it into c. Then,

$$h(x) = Ae^{\int f(x)dx},$$

where $A = e^c$. Now, observe that the exact values of c and A do not matter. If $h(x)$ is increased by a factor of A, the expression for y given by Eq. (3.33) does not change as the additional factors in the numerator and denominator cancel out. Hence, for the sake of convenience, c is usually set to zero. The expression for the integrating factor is then

$$h(x) = e^{\int f(x)dx}. \tag{3.34}$$

In practice, Eq. (3.33) is rarely used as it is often less taxing on your memory to find the integrating factor and go through the process of solving the differential equation.

Problem: Solve the following differential equation with the initial condition $y = 0$ when $x = 0$.

$$y' + xy = x.$$

The integrating factor is

$$h(x) = e^{\int x dx} = e^{\frac{x^2}{2}}.$$

Multiplying the integrating factor to both sides of the original differential equation,

$$\frac{d(e^{\frac{x^2}{2}} y)}{dx} = e^{\frac{x^2}{2}} x.$$

Separating variables and integrating,

$$e^{\frac{x^2}{2}} y = \int_0^x e^{\frac{x^2}{2}} x dx.$$

The expression on the right can be integrated using the substitution $u = \frac{x^2}{2}$ and $du = xdx$. Then,

$$\int_0^x e^{\frac{x^2}{2}} x dx = \int_0^{\frac{x^2}{2}} e^u du = [e^u]_0^{\frac{x^2}{2}} = e^{\frac{x^2}{2}} - 1$$

$$e^{\frac{x^2}{2}} y = e^{\frac{x^2}{2}} - 1.$$

Therefore, this differential equation has the solution

$$y = 1 - e^{-\frac{x^2}{2}}.$$

A Sneaky Case of an Integrating Factor

A rather common, but intimidating differential equation arises in problems where a system progressively gathers mass. It takes the form

$$y'' + f(y)y'^2 = g(y). \tag{3.35}$$

This equation can actually be simplified via an integrating factor. Since $y'' = \frac{1}{2}\frac{dy'^2}{dy}$,

$$\frac{dy'^2}{dy} + 2f(y)y'^2 = 2g(y).$$

This equation is of the same form as that in Eq. (3.32) with $y \to y'^2$ and $x \to y$ such that $y' \to \frac{dy'^2}{dy}$. Thus, the integrating factor is

$$e^{2\int f(y) dy}.$$

Then, one can multiply the above by the integrating factor, separate variables and simplify the differential equation above.

Problems

Vectors

1. *Lots of Practice*

(a) The magnitude of the addition of two vectors is four times the shorter one, and makes an angle of $60°$ with the shorter one. What is the ratio of the magnitudes of the vectors? What is the angle between them?

(b) Two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ subtend an angle θ. If $|\boldsymbol{a}| = 2|\boldsymbol{b}|$ and $|\boldsymbol{a} + \boldsymbol{b}| = \sqrt{3}|\boldsymbol{a} - \boldsymbol{b}|$, determine θ.

(c) If $|\boldsymbol{a} + \boldsymbol{b}| = |\boldsymbol{a} - \boldsymbol{b}|$, prove that $\boldsymbol{a} \perp \boldsymbol{b}$. Conversely, if $(\boldsymbol{a} + \boldsymbol{b}) \perp (\boldsymbol{a} - \boldsymbol{b})$, show that $|\boldsymbol{a}| = |\boldsymbol{b}|$.

(d) If $|\boldsymbol{a}| = 1$ and $|\boldsymbol{b}| = \frac{\sqrt{6}+\sqrt{2}}{2}$, and the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$ is $45°$, what is the angle between $\boldsymbol{a} + \boldsymbol{b}$ and $\boldsymbol{a}$?

(e) Consider two points located at position vectors $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, separated by a distance $r_{12} = |\boldsymbol{r}_1 - \boldsymbol{r}_2|$. Find a vector $\boldsymbol{A}$ from the origin to the point on the line between the two points at a distance $\frac{m}{m+n} r_{12}$ from the point at $\boldsymbol{r}_1$, where m and n are some real numbers. This result is known as the ratio theorem.

(f) In a methane molecule CH_4, each hydrogen atom is at the corner of a tetrahedron with the carbon atom at the centre. In a coordinate system centered about the carbon atom, if the direction of one of the C-H bonds is described by the vector $\boldsymbol{a} = \hat{\boldsymbol{i}} + \hat{\boldsymbol{j}} + \hat{\boldsymbol{k}}$ and the direction of an adjacent C-H bond is described by the vector $\boldsymbol{b} = \hat{\boldsymbol{i}} - \hat{\boldsymbol{j}} - \hat{\boldsymbol{k}}$, what is the angle θ between these two bonds?

(g) The dot product can be extended to n-dimensional vectors in $\mathbb{R}_n$ such that $\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$ is again the sum of the products of their corresponding components and $|\boldsymbol{a}|^2 = \boldsymbol{a} \cdot \boldsymbol{a}$. Prove the Cauchy-Schwartz inequality $|\boldsymbol{a} \cdot \boldsymbol{b}| \leq |\boldsymbol{a}| \cdot |\boldsymbol{b}|$. Hint: consider $|\boldsymbol{a}|\boldsymbol{b} - |\boldsymbol{b}|\boldsymbol{a}$ and $|\boldsymbol{a}|\boldsymbol{b} + |\boldsymbol{b}|\boldsymbol{a}$.

2. *Perpendicular Velocities**

Two particles move in a uniform region of downwards gravitational field g. They begin at the same location with initial velocities v_1 to the left and v_2 to the right, respectively. Find the distance between the two particles when their velocities are perpendicular to each other. Assume that they do not collide with each other nor with other entities.

3. *Halving Speed**

A particle moves with constant acceleration. At $t = 0$s, its speed is v. At $t = 1$s, its speed is $\frac{v}{2}$. At $t = 2$s, its speed is $\frac{v}{4}$. What is its speed at $t = 3$s?

4. *Sixth Collision***

An isolated system contains four particles that travel at constant velocities. Assuming that the particles pass through each other when they coincide and that a maximum of two particles can meet at a single point at every instance, there are a total of six possible intersections of particles. If you know that there are at least five encounters, can you predict if there will be a sixth (either at an earlier or later instance)?

Miscellaneous

5. *Catching Rain**

A stationary cylindrical bucket of radius r and height l is initially empty. It is currently raining and raindrops are evenly distributed about all space. The raindrops now travel vertically downwards at a constant speed v. Let the rate of volume of rain collected by the bucket now be r_1. Then, a horizontal wind now imparts the raindrops with a horizontal velocity u in the positive x-direction. Let the rate now be r_2. Compare r_1 and r_2.

6. *Running in the Rain**

Considering the second situation in the previous question, if the bucket is now replaced by a human running at a constant speed v_0 in the positive x-direction with the same dimensions as the cylinder, find the rate of volume of rain swept by the person, r. Note that the person can now "collect" rain sideways. Now, assuming that there is a shelter a distance d away, find the total volume of rain collected by the person as a function of v_0 (as he runs for shelter). Consider different cases. When is it not ideal to run faster?

7. *Smart Target**

A target takes the form of a two-dimensional square of side length l and can translate on the two-dimensional ground (but it cannot rotate). Supposing the center of the target can move a maximum distance l during the time that an arrow takes to land onto the ground, determine the minimum number of arrows that need to be fired simultaneously to guarantee a hit on the target.

Constant Acceleration and Projectile Motion

8. *Dropping Masses**

A mass with zero initial velocity is dropped from a height h_1 above the ground at time $t = 0$. At time $t = \tau$, another mass is dropped from rest from a height h_2. If the two masses first attain the same velocity at time $t = (k+1)\tau$ where k is a positive integer, show that the value

$$\sqrt{\frac{h_1}{h_2}}$$

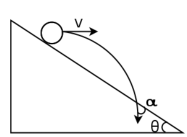
must be rational. The balls do not collide with each other. Assume that the balls are reflected at the same speed after colliding with the ground.

9. *Times Times**

A particle is thrown from a level ground at a certain fixed initial speed and variable elevation angle. There are two possible paths which result in the same range R. Let the times taken by the paths be t_1 and t_2 respectively. Find $t_1 t_2$.

10. *Ball on Inclined Plane**

A ball is thrown with velocity, v, horizontally off the surface of a plane with an angle of inclination θ as shown in the figure below. Find the angle α.



11. *Tossing over a Circle**

A vertical circular obstacle of radius R is placed on a horizontal ground. If a person is to toss a projectile at speed v such that its parabolic path is symmetrical about the vertical line through the center of circle and the projectile just touches the top of the circle, determine the maximum R for which the obstacle does not collide with the circle.

12. ***Maximum Range*****

A particle is thrown from an initial vertical coordinate y_0 at a fixed speed u and a variable angle θ with respect to the horizontal. If the landing ground is at $y = 0$, determine the maximum range and the angle θ that leads to this maximum range. Be careful about the case where $y_0 < 0$.

13. ***Maximum Range along Inclined Plane*****

A ball is tossed up an inclined plane with an angle of inclination θ, at an angle ϕ relative to the slope and fixed speed v. It then collides again with the plane at a later instant. Find ϕ that maximizes the distance between the initial position of the ball and the point of collision of the ball along the plane.

14. ***Throwing over a Thin Wall*****

A person wishes to throw a projectile over a wall on level ground. The projectile is thrown at coordinates $(0, 0)$ and is supposed to land at $(R, 0)$, $R > 0$. If there is a vertical wall with height h at x-coordinate $x_0 > 0$, what is the minimum initial speed at which the projectile must be thrown?

15. ***Watering a Garden******

An isotropic point source on the ground sprays water droplets at speed v uniformly in directions which make an angle between 0 and $\frac{\pi}{4}$ radians with the vertical. Define the origin to be at the point source and consider the plane of the ground. Suppose that the wetness at a point on the ground that is a distance r away from the origin is proportional to the number of water droplets impinging the immediate neighborhood of that point per unit area. Determine the radial coordinate r of a point that is $\frac{2}{(\sqrt{6}-\sqrt{2})}$ times as wet as the point that is a distance $\frac{v^2}{2g}$ away from the origin.

Polar Coordinates

16. ***Regular N-gon*****

Consider a system of $N \geq 3$ particles which are seemingly like the vertices of an imaginary regular polygon. The initial distance between adjacent particles is l. Number the particles from 1 to N in the clockwise direction. In the motion thereafter, the instantaneous velocity of the ith particle is always along the vector from the ith particle to the $(i+1)$th particle (the $(N+1)$th particle is taken to be the first particle). If the speeds of the particles are a constant v, determine the time when the distance between the particles becomes $\frac{l}{2}$. Furthermore, imagine an intangible polygon formed by connecting adjacent particles by intangible lines. Argue that the N particles must always be arranged in the shape of a regular N-gon, which is contracting. Furthermore, determine the angle that the imaginary N-gon has rotated about, by the time the adjacent distances become $\frac{l}{2}$.

17. *Constant Magnitude of Acceleration**

A particle wishes to travel in a circle of radius r and begins at rest. Supposing that its acceleration has a constant magnitude a, determine the distance travelled by the particle between the initial instance and the juncture when the particle attains its maximum angular velocity. Hint: In polar coordinates about the center of rotation, define α to be the instantaneous angle that the acceleration vector makes with the tangential direction. Then, differentiate a certain kinematic equation in polar coordinates. ("Introduction to Classical Mechanics")

18. *Searching in Fog***

In this two-dimensional problem, you are a ship operator and a pirate ship currently rests right next to you (the ships can be approximated as point particles). A dense fog suddenly emerges such that you are unable to locate the pirate ship. However, you know that the pirate ship will only travel at a constant velocity $\boldsymbol{v}$ in a particular direction. If you start your engine t_0 after the pirate ship begins moving, and travel at a speed u ($u > v$), what is the optimal path that you should take to guarantee catching the pirate in the minimum amount of time? What is this minimum amount of time? Assume that you can change the direction of your ship's velocity instantaneously.

Rigid Body Kinematics

19. *Two Velocities*

Two points on a rigid body are traveling at the same velocity $\boldsymbol{v}$. However, $\boldsymbol{v}$ is not perpendicular to the vector joining the two points. What can you say about the velocity of an arbitrary point on the rigid body?

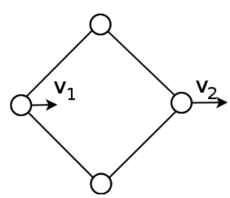
20. *Two Plates*

A rotating circular wheel of radius R is sandwiched between two parallel plates which are tangential to the wheel. The top plate travels at a velocity

v_1 while the bottom plate travels at a velocity v_2. Both velocities are along the x-direction. If there is no slipping between the wheel and the plates, determine the velocity of the center of the wheel and the wheel's angular velocity ω, which is defined to be positive anti-clockwise in the plane of the wheel.

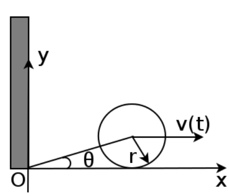
21. ***Pulling a Square****

Consider four identical rigids rods of length l connected by pinned connections in the form of a square. Suppose that you pull two diametrically opposite pins at velocities v_1 and v_2, respectively as shown in the figure below. Determine the velocity of the pin at the top. The entire motion lies in a single plane.



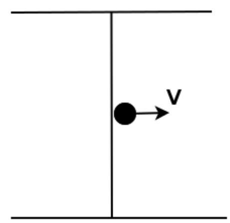
22. ***Moving Wheel****

Consider a wheel of radius r, rolling away from a wall with velocity $v(t)$ at its center. Define a coordinate system as shown in the figure below. Find the rate of change of the angular coordinate of the center of the wheel, $\dot{\theta}$, as a function of $v(t)$ and x, the x-coordinate of the center of the wheel with respect to the wall.



23. ***H-Shape*****

Consider 6 identical rods of length l that are connected by fixed connections to form a rigid "H"-shape as shown in the figure on the next page. At the current instance, a particle is right next to the center of mass of the H-shape as depicted. If the constant velocity and angular velocity of the H-shape are v_{CM} rightwards and ω clockwise respectively, determine the minimum rightwards velocity of the particle v, such that it can escape without colliding with the structure.



Differential Equations

24. *Inverse-Squared Force**

Suppose you obtain the equation of motion $\ddot{r} = \frac{k}{r^2}$. Solve for the particle's velocity $\dot{r}$ as a function of r.

25. *Bernoulli's Equation***

Consider a differential equation of the form

$$y' + f(x)y = g(x)y^n.$$

Show that the substitution $z = y^{1-n}$ yields an equation that can be simplified by an integrating factor. Determine the integrating factor.

26. *Bug on Rubber Band****

A bug is initially located at the fixed, left end of a rubber band of initial length l. The right end is then pulled away at a velocity v. If the rubber band stretches uniformly and the bug travels at a velocity u relative to the rubber band, determine the time taken for the bug to reach the other end of the rubber band. (BAUPC)

Solutions

1. *Lots of Practice*

(a) Let the two vectors be $\boldsymbol{a}$ and $\boldsymbol{b}$ with $|\boldsymbol{b}| < |\boldsymbol{a}|$ and define θ as the angle subtended by the two vectors.

$$|\boldsymbol{a}+\boldsymbol{b}| = 4|\boldsymbol{b}| \implies a^2 + b^2 + 2\boldsymbol{a}\cdot\boldsymbol{b} = 16b^2$$

$$(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{b} = |\boldsymbol{a}+\boldsymbol{b}|\cdot|\boldsymbol{b}|\cos 60^\circ$$

$$\implies \boldsymbol{a}\cdot\boldsymbol{b} + b^2 = 2b^2$$

as $|\boldsymbol{a}+\boldsymbol{b}| = 4|\boldsymbol{b}|$. Shifting b^2 to the right-hand side,

$$\boldsymbol{a}\cdot\boldsymbol{b} = b^2.$$

Substituting this expression into the first equation,

$$a^2 = 13b^2$$

$$|\boldsymbol{a}| = \sqrt{13}|\boldsymbol{b}|$$

$$\theta = \cos^{-1}\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|} = \cos^{-1}\frac{1}{\sqrt{13}}.$$

(b) Since $|\boldsymbol{a}+\boldsymbol{b}| = \sqrt{3}|\boldsymbol{a}-\boldsymbol{b}|$,

$$(\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}+\boldsymbol{b}) = 3(\boldsymbol{a}-\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b})$$

$$8\boldsymbol{a}\cdot\boldsymbol{b} = 2(a^2+b^2) = 6b^2$$

$$\boldsymbol{a}\cdot\boldsymbol{b} = \frac{3}{4}b^2$$

$$\theta = \cos^{-1}\frac{\boldsymbol{a}\cdot\boldsymbol{b}}{|\boldsymbol{a}||\boldsymbol{b}|} = \cos^{-1}\frac{3}{8}.$$

(c)

$$|\boldsymbol{a}+\boldsymbol{b}| = |\boldsymbol{a}-\boldsymbol{b}|$$

$$\implies a^2 + 2\boldsymbol{a}\cdot\boldsymbol{b} + b^2 = a^2 - 2\boldsymbol{a}\cdot\boldsymbol{b} + b^2$$

$$\implies \boldsymbol{a}\cdot\boldsymbol{b} = 0,$$

which shows that they are perpendicular. Moving on to the next part, since $(\boldsymbol{a}+\boldsymbol{b}) \perp (\boldsymbol{a}-\boldsymbol{b})$,

$$(\boldsymbol{a}+\boldsymbol{b})\cdot(\boldsymbol{a}-\boldsymbol{b}) = 0$$

$$\implies a^2 = b^2$$
$$\implies |\boldsymbol{a}| = |\boldsymbol{b}|.$$

(d) Let θ denote the angle between $\boldsymbol{a}+\boldsymbol{b}$ and $\boldsymbol{a}$.

$$\begin{aligned}
\theta &= \cos^{-1}\frac{(\boldsymbol{a}+\boldsymbol{b})\cdot\boldsymbol{a}}{|\boldsymbol{a}+\boldsymbol{b}||\boldsymbol{a}|} \\
&= \cos^{-1}\frac{a^2+\boldsymbol{a}\cdot\boldsymbol{b}}{\sqrt{a^2+b^2+2\boldsymbol{a}\cdot\boldsymbol{b}}\cdot|\boldsymbol{a}|} \\
&= \cos^{-1}\frac{1^2+1\times\frac{\sqrt{6}+\sqrt{2}}{2}\times\cos 45^\circ}{\sqrt{1^2+\left(\frac{\sqrt{6}+\sqrt{2}}{2}\right)^2+2\times 1\times\frac{\sqrt{6}+\sqrt{2}}{2}\times\cos 45^\circ}\cdot 1} \\
&= \cos^{-1}\frac{1+\frac{\sqrt{3}+1}{2}}{\sqrt{2\sqrt{3}+4}} = \cos^{-1}\frac{\frac{\sqrt{3}+3}{2}}{\sqrt{3}+1} \\
&= \cos^{-1}\frac{\sqrt{3}}{2} \\
&= 30^\circ.
\end{aligned}$$

(e)

$$\boldsymbol{A} = \boldsymbol{r}_1 + \frac{m}{m+n}(\boldsymbol{r}_2-\boldsymbol{r}_1) = \frac{m\boldsymbol{r}_2+n\boldsymbol{r}_1}{m+n}.$$

(f)

$$\theta = \cos^{-1}\frac{\begin{pmatrix}1\\1\\1\end{pmatrix}\cdot\begin{pmatrix}1\\-1\\-1\end{pmatrix}}{|\boldsymbol{a}||\boldsymbol{b}|} = \cos^{-1}\frac{1-1-1}{3} = \cos^{-1}-\frac{1}{3} = 109.5^\circ\ (4\text{s.f.}).$$

(g) If $\boldsymbol{a}$ or $\boldsymbol{b}$ is the null vector, the equality case of the Cauchy-Schwartz inequality evidently holds so we only need to consider the non-trivial cases. Firstly, observe that the dot product of any arbitrary vector $\boldsymbol{A}$ with itself must be non-negative as

$$A_1^2 + A_2^2 + \cdots + A_n^2 \geq 0.$$

The equality case only occurs when $\boldsymbol{A} = \mathbf{0}$ is the null vector. Applying this inequality to $|\boldsymbol{a}|\boldsymbol{b} - |\boldsymbol{b}|\boldsymbol{a}$,

$$\begin{aligned}(|\boldsymbol{a}|\boldsymbol{b} - |\boldsymbol{b}|\boldsymbol{a}) \cdot (|\boldsymbol{a}|\boldsymbol{b} - |\boldsymbol{b}|\boldsymbol{a}) &\geq 0 \\ 2a^2b^2 - 2|\boldsymbol{a}||\boldsymbol{b}|\boldsymbol{a} \cdot \boldsymbol{b} &\geq 0 \\ a^2b^2 &\geq |\boldsymbol{a}||\boldsymbol{b}|\boldsymbol{a} \cdot \boldsymbol{b} \\ \implies |\boldsymbol{a}||\boldsymbol{b}| &\geq \boldsymbol{a} \cdot \boldsymbol{b}.\end{aligned}$$

Repeating this process with $|\boldsymbol{a}|\boldsymbol{b} + |\boldsymbol{b}|\boldsymbol{a}$, one can also show that

$$|\boldsymbol{a}||\boldsymbol{b}| \geq -\boldsymbol{a} \cdot \boldsymbol{b},$$

which shows that

$$|\boldsymbol{a}||\boldsymbol{b}| \geq |\boldsymbol{a} \cdot \boldsymbol{b}|.$$

Evidently, the equality case only occurs when $|\boldsymbol{a}|\boldsymbol{b} - |\boldsymbol{b}|\boldsymbol{a} = \mathbf{0}$ or $|\boldsymbol{a}|\boldsymbol{b} + |\boldsymbol{b}|\boldsymbol{a} = \mathbf{0}$. That is, the two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ must either be parallel or anti-parallel.

2. *Perpendicular Velocities**

Define the x and y-axes to be positive rightwards and upwards. The velocities of the particles at time t are respectively $(-v_1, -gt)$ and $(v_2, -gt)$. For these to be perpendicular,

$$\begin{aligned}\begin{pmatrix}-v_1 \\ -gt\end{pmatrix} \cdot \begin{pmatrix}v_2 \\ -gt\end{pmatrix} &= 0 \\ \implies v_1v_2 &= g^2t^2 \\ t &= \frac{\sqrt{v_1v_2}}{g}.\end{aligned}$$

The distance between the particles is simply their difference in x-coordinates which is

$$(v_1 + v_2)t = \frac{(v_1 + v_2)\sqrt{v_1v_2}}{g}.$$

3. *Halving Speed**

Let the initial velocity be $\boldsymbol{v}$ and the constant acceleration be $\boldsymbol{a}$. The conditions of the question imply

$$|\boldsymbol{v}+\boldsymbol{a}| = \frac{v}{2} \implies 2\boldsymbol{a}\cdot\boldsymbol{v} + a^2 = -\frac{3}{4}v^2$$

$$|\boldsymbol{v}+2\boldsymbol{a}| = \frac{v}{4} \implies 4\boldsymbol{a}\cdot\boldsymbol{v} + 4a^2 = -\frac{15}{16}v^2.$$

Solving these equations simultaneously,

$$\boldsymbol{a}\cdot\boldsymbol{v} = -\frac{33}{64}v^2$$

$$a^2 = \frac{9}{32}v^2$$

$$\implies |\boldsymbol{v}+3\boldsymbol{a}| = \sqrt{v^2 + 6\boldsymbol{a}\cdot\boldsymbol{v} + 9a^2} = \frac{\sqrt{7}}{4}v.$$

4. *Sixth Collision***

Without the loss of generality, number the particles from 1 to 4 and suppose that only particles 1 and 2 have yet to collide. Defining $\boldsymbol{r}_i$ and $\boldsymbol{v}_i$ as the initial position vector and velocity of the ith particle and t_{ij} to be the time of the collision between the ith and jth particle, we have

$$\boldsymbol{r}_1 - \boldsymbol{r}_3 = (\boldsymbol{v}_3 - \boldsymbol{v}_1)t_{13},$$

$$\boldsymbol{r}_1 - \boldsymbol{r}_4 = (\boldsymbol{v}_4 - \boldsymbol{v}_1)t_{14},$$

$$\boldsymbol{r}_2 - \boldsymbol{r}_3 = (\boldsymbol{v}_3 - \boldsymbol{v}_2)t_{23},$$

$$\boldsymbol{r}_2 - \boldsymbol{r}_4 = (\boldsymbol{v}_4 - \boldsymbol{v}_2)t_{24},$$

$$\boldsymbol{r}_3 - \boldsymbol{r}_4 = (\boldsymbol{v}_4 - \boldsymbol{v}_3)t_{34},$$

as these pairs of particles collide. Subtracting the first equation from the second,

$$\boldsymbol{r}_3 - \boldsymbol{r}_4 = t_{14}\boldsymbol{v}_4 - t_{13}\boldsymbol{v}_3 + (t_{13} - t_{14})\boldsymbol{v}_1 = t_{14}(\boldsymbol{v}_4 - \boldsymbol{v}_3) + (t_{13} - t_{14})(\boldsymbol{v}_1 - \boldsymbol{v}_3).$$

Equating this with the fifth equation,

$$(t_{13} - t_{14})(\boldsymbol{v}_1 - \boldsymbol{v}_3) = (t_{34} - t_{14})(\boldsymbol{v}_4 - \boldsymbol{v}_3).$$

Applying a similar process to the third and fourth equation, we obtain

$$(t_{23} - t_{24})(\boldsymbol{v}_2 - \boldsymbol{v}_3) = (t_{34} - t_{24})(\boldsymbol{v}_4 - \boldsymbol{v}_3).$$

The last two equations show that

$$\boldsymbol{v}_1 - \boldsymbol{v}_2 = \left(\frac{t_{34} - t_{14}}{t_{13} - t_{14}} - \frac{t_{34} - t_{24}}{t_{23} - t_{24}}\right)(\boldsymbol{v}_4 - \boldsymbol{v}_3),$$

where we have used the fact that $t_{13} \neq t_{14}$ and $t_{23} \neq t_{24}$ as a maximum of two particles can coincide at a single location at every juncture.

$$\begin{aligned}\boldsymbol{r}_1 - \boldsymbol{r}_2 = (\boldsymbol{r}_1 - \boldsymbol{r}_3) - (\boldsymbol{r}_2 - \boldsymbol{r}_3) = (\boldsymbol{v}_3 - \boldsymbol{v}_1)t_{13} - (\boldsymbol{v}_3 - \boldsymbol{v}_2)t_{23} \\ = \left(\frac{t_{34} - t_{14}}{t_{14} - t_{13}}t_{13} - \frac{t_{34} - t_{24}}{t_{24} - t_{23}}t_{23}\right)(\boldsymbol{v}_4 - \boldsymbol{v}_3).\end{aligned}$$

Therefore,

$$\boldsymbol{r}_1 - \boldsymbol{r}_2 = \frac{\frac{t_{34} - t_{14}}{t_{13} - t_{14}}t_{13} - \frac{t_{34} - t_{24}}{t_{23} - t_{24}}t_{23}}{\frac{t_{34} - t_{14}}{t_{13} - t_{14}} - \frac{t_{34} - t_{24}}{t_{23} - t_{24}}}(\boldsymbol{v}_2 - \boldsymbol{v}_1),$$

which shows that particles 1 and 2 collide. A more intuitive solution would be to observe that there exists a particle X which is guaranteed to collide with the other three particles. Consider the frame of X and define the origin at X. Since there are at least five collisions, there are at least 2 pairs of other particles (of which all three other particles are included at least once each) which intersect. Since these three particles must also pass through the origin, their trajectories must be along a single straight line through the origin (the other alternative is for them to pass through the origin at the same time but this is forbidden by the fact that a maximum of two particles can cross at any location) — implying that all 6 pairs must collide.

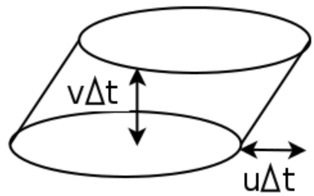
5. *Catching Rain**

In both situations, consider the amount of rain collected by the bucket in a time interval Δt. In the first situation, the volume collected is a cylinder of radius r and length $v\Delta t$. Hence, the additional volume collected in a time interval Δt is

$$\Delta V = \pi r^2 v \Delta t$$

$$r_1 = \frac{\Delta V}{\Delta t} = \pi r^2 v.$$

In the second scenario, the volume collected is of the following shape



The volume of this object is still the area of the base multiplied by the perpendicular height.

$$\Delta V = \pi r^2 \cdot h = \pi r^2 v \Delta t$$

$$r_2 = \frac{\Delta V}{\Delta t} = \pi r^2 v.$$

This becomes obvious when we consider the fact that a flat surface with an area vector $\boldsymbol{A}$ sweeps out volume in a medium at a rate $\boldsymbol{A} \cdot \boldsymbol{v}$ where $\boldsymbol{v}$ is its relative velocity with the medium. Hence,

$$r_1 = r_2.$$

6. *Running in the Rain**

The rate r has two contributions — due to the rain pouring on his head, similar to the situation with bucket, and due to the additional rain he runs into sideways. Let these two rates be r_1 and r_2 respectively. Then,

$$r = r_1 + r_2.$$

r_1 has been computed in the previous question as

$$r_1 = \pi r^2 v.$$

r_2 is simply the rate at which raindrops are swept up by the side of the human. The relative horizontal velocity between the human and the raindrops is $v_0 - u$. Hence, in time Δt, the volume of rain colliding with the side of the human is

$$\Delta V = 2lr|v_0 - u|\Delta t,$$

where $2lr$ is the cross-sectional area of the cylinder. Note that the absolute value sign is necessary as the raindrops may impinge on the human from both sides.

$$r_2 = 2lr|v_0 - u|.$$

The total rate is

$$r = \pi r^2 v + 2lr|v_0 - u|.$$

The person takes a time $t = \frac{d}{v_0}$ to reach the shelter. Hence, the total volume of rain swept by the person is

$$V = r \cdot t = \frac{\pi r^2 vd + 2lr|v_0 - u|d}{v_0}.$$

Next, a few cases must be considered. When $v_0 \geq u$, $|v_0 - u| = v_0 - u$.

$$V = 2lrd + \frac{dr}{v_0}(\pi rv - 2lu).$$

If $v > \frac{2l}{\pi r}u$, the term in the bracket is positive and the total volume swept, V, decreases as v_0 increases. It is ideal to travel at a large velocity to minimize the amount of time spent in the rain. However, if $v < \frac{2l}{\pi r}u$, V increases as v_0 increases. In this situation, it is better to reduce v_0. However, v_0 must still be greater than u. Hence, the minimum value amount of rain swept in this regime occurs when $v_0 = u$. Then,

$$V = \frac{d\pi r^2 v}{u}.$$

In this situation, the person is traveling at the same horizontal velocity as the raindrops such that no rain hits the person from the side. The only rain gathered by the person is that pouring onto his head.

For the case where $v_0 \leq u$, $|v_0 - u| = u - v_0$.

$$V = \frac{dr}{v_0}(\pi rv + 2lu) - 2lrd.$$

Hence, it is still better to run faster in this case. The maximum value of v_0 in this regime is also u. Then, the total volume of rain collected is

$$V = \frac{d\pi r^2 v}{u}.$$

7. *Smart Target**

The final position of the target's center can be anywhere within a circle of radius l, centered about the initial location of its center. Observe that an arrow shot at a particular point P eliminates the possibility of the center of mass falling within a square of side length l, centered at P. Thus, we just need to cover a circle of radius l with squares of length l. It is easy to see that this is possible with 4 squares. We can also prove that this is impossible with three squares by considering their areas. The area of three squares is $3l^2$ which is smaller than the area of the circle πl^2. Thus, four arrows are necessary to guarantee a hit.

8. *Dropping Masses**

Define the positive direction to be upwards. The crucial observation is that the two balls first attain the same velocity when the ball with the smaller maximum speed attains its maximum positive velocity. To see why this is so, suppose that we observe the two balls to have the same velocity at the current instance. Then, we can rewind their motions slightly and observe their velocities at an earlier instance — their velocities will still be identical unless one ball has just rebounded from the ground. Furthermore, this reflected ball must be the ball with the smaller maximum speed as it would take less time for it to decelerate from its maximum positive velocity to its current velocity. Therefore, the earliest juncture at which the two balls possess the same velocity is when the ball with the smaller maximum speed has just rebounded from the ground and when the other ball simultaneously attains this velocity. As such, we have to consider two cases — when $h_1 \geq h_2$ and $h_1 < h_2$. In the case of the former, the balls first satisfy the required condition when the second ball just rebounds from the ground. Observing that the "periods" of the balls' velocities are $\sqrt{\frac{2h_1}{g}}$ and $\sqrt{\frac{2h_2}{g}}$ respectively,

$$k\tau = (2m+1)\sqrt{\frac{2h_2}{g}},$$

where m is an integer for the second ball to just rebound from the ground. Furthermore,

$$(k+1)\tau = 2\sqrt{\frac{2h_1}{g}} - \sqrt{\frac{2h_2}{g}} + 2n\sqrt{\frac{2h_1}{g}} = (2n+2)\sqrt{\frac{2h_1}{g}} - \sqrt{\frac{2h_2}{g}},$$

where n is an integer for the first ball to attain the velocity $\sqrt{2gh_2}$. Note that $2\sqrt{\frac{2h_1}{g}} - \sqrt{\frac{2h_2}{g}}$ is the time that the first ball takes to first attain this velocity. Eliminating the τ's,

$$\sqrt{\frac{h_1}{h_2}} = \frac{(k+1)(2m+1)+k}{k(2n+2)}.$$

When $h_2 > h_1$, one can similarly show that the required conditions are

$$(k+1)\tau = 2n\sqrt{\frac{2h_1}{g}}$$

$$k\tau = (2m+2)\sqrt{\frac{2h_2}{g}} - \sqrt{\frac{2h_1}{g}}$$

$$\implies \sqrt{\frac{h_1}{h_2}} = \frac{(k+1)(2m+2)}{2kn+k+1}.$$

9. *Times Times**

Define the origin at the launching point. The vertical displacement at time t is

$$s_y = v_y t - \frac{1}{2}gt^2,$$

where v_y is the initial vertical velocity. The non-trivial solution to $s_y = 0$ is

$$t = \frac{2v_y}{g}.$$

Another way to obtain this answer would be to compute the time the particle takes to reach the peak and then multiply it by two due to the symmetrical property of projectile motion. Moving on, the range is simply the horizontal distance that the projectile has travelled at this juncture.

$$R = v_x t = \frac{2v_x v_y}{g}.$$

Expressing $v_x = v\cos\theta$ and $v_y \sin\theta$ in terms of the speed v and elevation angle θ,

$$R = \frac{v^2 \sin 2\theta}{g}.$$

Evidently, the other angle is $\frac{\pi}{2} - \theta$ as $\sin(\pi - x) = \sin x$. The two times taken are then

$$t_1 = \frac{2v\sin\theta}{g}$$

$$t_2 = \frac{2v\cos\theta}{g}$$

$$t_1 t_2 = \frac{2v^2\sin 2\theta}{g^2} = \frac{2R}{g}.$$

Another way is to first express $\sin\theta$ and $\cos\theta$ via

$$t = \frac{2v\sin\theta}{g} \implies \sin\theta = \frac{gt}{2v}$$

$$R = v\cos\theta t \implies \cos\theta = \frac{R}{vt},$$

where we have eliminated the possibility of $t = 0$. Squaring the trigonometric terms and adding them together,

$$\frac{g^2t^2}{4v^2} + \frac{R^2}{v^2t^2} = 1$$

$$t^4 - \frac{4v^2}{g^2}t^2 + \frac{4R^2}{g^2} = 0.$$

This is a quadratic equation in t^2 and should only have two solutions t_1^2 and t_2^2 (the squares of the two times of interest). By Vieta's theorem, we know that the product of the roots of a quadratic equation yields

$$t_1^2t_2^2 = \frac{4R^2}{g^2}$$

$$t_1t_2 = \frac{2R}{g},$$

where the times must be positive.

10. *Ball on Inclined Plane**

Define the x and y-axes to be positive rightwards and downwards respectively. From the kinematics equations,

$$s_y = \frac{1}{2}gt^2,$$

$$s_x = vt.$$

When the ball collides with the plane, $s_y = s_x\tan\theta$.

$$t = \frac{2v\tan\theta}{g}.$$

Thus, the vertical and horizontal velocities at this juncture are

$$v_y = gt = 2v\tan\theta,$$

$$v_x = v.$$

The components of the ball's velocity parallel and perpendicular to the plane are

$$v_{\parallel} = v_y \sin\theta + v_x \cos\theta = v\left(\frac{2\sin^2\theta}{\cos\theta} + \cos\theta\right),$$

$$v_{\perp} = v_y \cos\theta - v_x \sin\theta = v\sin\theta,$$

$$\alpha = \tan^{-1}\frac{v_{\perp}}{v_{\parallel}} = \tan^{-1}\frac{\sin\theta\cos\theta}{1+\sin^2\theta}.$$

11. *Tossing over a Circle**

The vertical velocity of the projectile must be zero at the top of the circle. This implies that the vertical component of the initial velocity of the projectile obeys

$$v_y^2 = 4gR.$$

The horizontal component of the squared velocity, which remains constant throughout the motion, is then

$$v_x^2 = v^2 - 4gR.$$

Since the path of a projectile is reversible, the projectile will not collide with the circle if and only if the same projectile tossed off the top of a circle with only a horizontal velocity v_x (and no vertical velocity) is found to not collide with the circle too. The latter condition is satisfied if the distance between the projectile and the center of the circle is greater than R at all times.

$$(v_x t)^2 + \left(R - \frac{1}{2}gt^2\right)^2 > R^2,$$

where t is measured from the juncture where the particle is at the top of the circle. Simplifying,

$$\frac{1}{4}g^2t^2 + (v^2 - 5gR) > 0.$$

We obtain the desired result if $R < \frac{v^2}{5g}$. Another way of seeing this is to require the centripetal acceleration at the top of the circle to be greater than g (so that the particle instantaneously travels along an arc with a larger radius of curvature).

$$\frac{v_x^2}{R} > g$$
$$\implies \frac{v^2}{R} > 5g$$
$$R < \frac{v^2}{5g}.$$

12. *Maximum Range***

The trajectory of the particle is described by the parabola:

$$y = y_0 + \tan\theta x - \frac{g}{2u^2}\sec^2\theta x^2.$$

Let the maximum range be R. When $x = R$, the trajectory equation satisfies

$$y_0 + \tan\theta R - \frac{g}{2u^2}\sec^2\theta R^2 = 0.$$

Well, we could solve for R to get

$$R = \frac{\tan\theta + \sqrt{\tan^2\theta + \frac{2gy_0}{u^2}\sec^2\theta}}{\frac{g}{u^2}\sec^2\theta}. \tag{3.36}$$

This is an important equation that determines the range of a projectile in terms of u and θ (we have rejected the other solution which is smaller). We could continue to differentiate it with respect to θ and solve for $\frac{dR}{d\theta} = 0$ but there is a slick method. Implicitly differentiating the previous equation with respect to θ and using the fact that $\frac{dR}{d\theta} = 0$ when R is a stationary point,

$$\sec^2\theta R - \frac{g}{u^2}\sec^2\theta\tan\theta R^2 = 0$$
$$R = \frac{u^2}{g}\cot\theta,$$

as we reject the trivial solution $R = 0$. Technically, we can check that this is indeed a maximum point by finding the second derivative of R with respect to θ via Eq. (3.36) but we shall avoid this tedium. When R is a maximum,

$$\tan\theta = \frac{u^2}{gR}.$$

This expression for $\tan\theta$ can be substituted into the trajectory equation to obtain

$$y_0 + \frac{u^2}{g} - \frac{g}{2u^2}\left(\frac{u^4}{g^2R^2} + 1\right)R^2 = 0,$$

where we have used the fact that $\sec^2\theta = \tan^2\theta + 1$. Then,

$$R = \frac{u}{g}\sqrt{u^2 + 2gy_0},$$

and

$$\begin{aligned}\tan\theta &= \frac{u^2}{gR} = \frac{u}{\sqrt{u^2 + 2gy_0}}\\ \theta &= \tan^{-1}\frac{u}{\sqrt{u^2 + 2gy_0}}.\end{aligned}$$

Lastly, there is another slight technicality. In the case of $y_0 < 0$, how do we know which x-coordinate this R–which is a maximum point–corresponds to (as there are two x-coordinates where $y = 0$)? That is, is it really the range of the projectile? Well, one can solve the quadratic trajectory equation for two values of x — substituting the above expression for $\tan\theta$ will show that R indeed corresponds to the larger root and is hence the range. Finally, this range assumes that the particle is able to reach that larger x-coordinate without facing any blockages.

A direct corollary of the above result is that the maximum range in the case of a level ground ($y_0 = 0$) is

$$R = \frac{u^2}{g},$$

and occurs when $\theta = \frac{\pi}{4}$ radians.

13. *Maximum Range along Inclined Plane***

Define the x-axis to be parallel to the slope, pointing upwards, and the y-axis to be normal to the slope, pointing away from the slope. Then, the ball experiences constant accelerations $-g\sin\theta$ and $-g\cos\theta$ in the x and y directions, respectively. The coordinates of the ball at a time t are given by

$$y = v\sin\phi t - \frac{1}{2}g\cos\theta t^2,$$

$$x = v\cos\phi t - \frac{1}{2}g\sin\theta t^2.$$

When $y = 0$, the non-trivial solution is

$$t = \frac{2v\sin\phi}{g\cos\theta}.$$

At this juncture,

$$x = \frac{2v^2}{g\cos\theta}(\sin\phi\cos\phi - \sin^2\phi\tan\theta) = \frac{2v^2}{g\cos^2\theta}\sin\phi(\cos\phi\cos\theta - \sin\phi\sin\theta).$$

Applying the trigonometric identities $\cos(A+B) = \cos A\cos B - \sin A\sin B$ and $2\sin A\cos B = \sin(A+B) + \sin(A-B)$,

$$x = \frac{2v^2}{g\cos^2\theta}\sin\phi\cos(\phi+\theta) = \frac{v^2}{g\cos^2\theta}\left[\sin(2\phi+\theta) - \sin\theta\right],$$

which is evidently maximum when

$$2\phi + \theta = \frac{\pi}{2}$$

$$\phi = \frac{\pi}{4} - \frac{\theta}{2}.$$

Other values of ϕ — namely the above expression plus an integer multiple of $\frac{\pi}{2}$ — are infeasible as ϕ and θ must obviously be acute and larger than zero. When $\phi = \frac{\pi}{4} - \frac{\theta}{2}$,

$$x = \frac{v^2(1-\sin\theta)}{g\cos^2\theta}.$$

14. *Throwing over a Thin Wall***

Note that since the starting and ending points are fixed and three points define a parabola of the form $y = ax^2 + bx + c$, if we choose a particular y-coordinate of the projectile at x_0, we will have fixed the parabola and hence will be able to determine the initial speed and angle the projectile was thrown at. Furthermore, we know that $\theta = \frac{\pi}{4}$ should correspond to the smallest initial velocity if the particle along this trajectory can pass over the wall. Hence, we must determine if this is so.

From the trajectory equation,

$$y = \tan\theta x - \frac{g\sec^2\theta}{2u^2}x^2.$$

The range on a level ground is obtained when $y = 0$.

$$R = \frac{2u^2\sin\theta\cos\theta}{g}$$

$$\implies u^2 = \frac{Rg}{2\sin\theta\cos\theta}.$$

Substituting this into the trajectory equation,

$$\tan\theta x - \frac{\tan\theta}{R}x^2 = y.$$

If we let $y(x_0) = y_0$,

$$\tan\theta = \frac{y_0}{x_0 - \frac{x_0^2}{R}}.$$

Note that the denominator is strictly greater than zero as $0 < x_0 < R$. If we let the denominator be k where $k > 0$,

$$\sec^2\theta = \frac{y_0^2}{k^2} + 1$$

$$\cos^2\theta = \frac{k^2}{y_0^2 + k^2}$$

$$\cos\theta = \frac{k}{\sqrt{y_0^2 + k^2}}.$$

We reject negative values of $\cos\theta$ as $0 < \theta < \frac{\pi}{2}$. Similarly,

$$\sin\theta = \frac{y_0}{\sqrt{y_0^2 + k^2}}.$$

Then,

$$u^2 = \frac{Rg}{2\sin\theta\cos\theta} = \frac{Rg(y_0^2 + k^2)}{2y_0k} = \frac{Ry_0}{2k} + \frac{Rk}{2y_0}.$$

To minimize u^2, we find its first derivative with respect to y_0.

$$\frac{d(u^2)}{dy_0} = \frac{Rg}{2k} - \frac{Rkg}{2y_0^2}.$$

The stationary point occurs when

$$y_0 = k = x_0 - \frac{x_0^2}{R},$$

where we have rejected the negative solution as the ground is at $y = 0$. Then

$$\cos\theta = \frac{1}{\sqrt{2}}$$

$$\implies \theta = \frac{\pi}{4}$$

$$u = \sqrt{\frac{gR}{2}}.$$

Hence, the minimum speed that the projectile should be thrown at is $\sqrt{\frac{gR}{2}}$, at an angle $\theta = \frac{\pi}{4}$ when $y_0 = x_0 - \frac{x_0^2}{R} \geq h$. (i.e. the projectile is above h at x_0 when $\theta = \frac{\pi}{4}$). This is as expected. In the case where $h > x_0 - \frac{x_0^2}{R}$, the minimum speed occurs when $y_0 = h$ as

$$\frac{d(u^2)}{dy_0} = \frac{Rg}{2k} - \frac{Rkg}{2y_0^2} > 0,$$

for all $y_0 > x_0 - \frac{x_0^2}{R}$. Then, substituting $y_0 = h$ into the expression for u^2, the minimum speed is

$$u = \sqrt{\frac{Rg(h^2+k^2)}{2hk}},$$

and occurs when

$$\tan\theta = \frac{h}{k},$$

where $k = x_0 - \frac{x_0^2}{R}$.

15. *Watering a Garden****

Consider a sphere that represents the possible directions of the velocities of the water droplets.

Let θ be the angle between the velocity of a water droplet and the vertical axis. Then, we can ascribe all water droplets a θ coordinate from 0 to $\frac{\pi}{4}$ radians. Notice that the amount of water droplets between angles θ and $\theta + d\theta$ is proportional to $2\pi \sin\theta d\theta$ (proportional to the curved surface area of the disk in Fig. 3.28). Suppose that the final radial coordinate of the water droplets between angles θ and $\theta + d\theta$, after they have fallen to the ground,

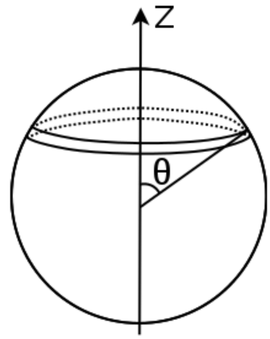


Figure 3.28: Area between θ and $\theta + d\theta$

lies between r and $r + dr$. The area on the ground that they occupy is then $2\pi r dr$. Thus, the wetness at a radial coordinate r is proportional to $\frac{\sin\theta d\theta}{rdr}$. Our objective is to determine this function. It can be easily proven that the range of a droplet at angle θ is

$$r = \frac{v^2 \sin 2\theta}{g}.$$

Evidently, $r = \frac{v^2}{2g}$ corresponds to $\theta = \frac{\pi}{12}$. Furthermore,

$$\frac{dr}{d\theta} = \frac{2v^2 \cos 2\theta}{g}$$

$$\frac{rdr}{\sin\theta d\theta} = \frac{2v^4 \sin 2\theta \cos 2\theta}{g^2 \sin\theta} = \frac{4v^4}{g^2} \cos\theta \cos 2\theta.$$

This is inversely proportional to the wetness. Thus, our desired θ obeys

$$\cos\theta \cos 2\theta = \frac{(\sqrt{6} - \sqrt{2})}{2} \cos\frac{\pi}{12} \cos\frac{\pi}{6} = \frac{\sqrt{3}}{4},$$

where we have used the fact that $\cos\frac{\pi}{12} = \frac{\sqrt{6}+\sqrt{2}}{4}$. Expressing the left-hand side in terms of $\cos\theta$ only,

$$2\cos^3\theta - \cos\theta - \frac{\sqrt{3}}{4} = 0$$

$$\left(\cos\theta - \frac{\sqrt{3}}{2}\right)\left(2\cos^2\theta - \sqrt{3}\cos\theta + \frac{1}{2}\right) = 0.$$

Therefore, the only solution is $\theta = \frac{\pi}{6}$ as the discriminant of the expression in the second bracket is negative. The corresponding radial coordinate is

$$r = \frac{\sqrt{3}v^2}{2g}.$$

16. *Regular N-gon***

Let $r(t)$ be the instantaneous distance between adjacent particles. The relative radial velocity between adjacent particles is

$$\dot{r} = v\left(\cos\frac{2\pi}{N} - 1\right).$$

Then,

$$r = l + v\left(\cos\frac{2\pi}{N} - 1\right)t.$$

When $r = \frac{l}{2}$,

$$t = \frac{l}{2v\left(1 - \cos\frac{2\pi}{N}\right)}.$$

To argue that the particles maintain the shape of a regular N-gon, notice that all adjacent distances must be r by symmetry — implying that they form a regular N-gon. To calculate the angle that this imaginary N-gon has rotated, we observe that the center (more technically, centroid) must be the center of rotation. We then calculate the angular velocity of the imaginary N-gon. The distance between the center and a particle can be proved by geometrical means to be

$$\frac{r}{2\sin\frac{\pi}{N}}.$$

The component of velocity tangential to the line joining the center and a particle is

$$v\cos\frac{\pi}{N}.$$

Hence, the angular velocity is

$$\omega = \frac{v\cos\frac{\pi}{N}}{\frac{r}{2\sin\frac{\pi}{N}}} = \frac{v}{r}\sin\frac{2\pi}{N} = \frac{v}{l + v\left(\cos\frac{2\pi}{N} - 1\right)t}\sin\frac{2\pi}{N}.$$

The total angle that it has rotated between $t = 0$ and $t = \frac{l}{2v(1-\cos\frac{2\pi}{N})}$ is then

$$\Delta\theta = \frac{v\sin\frac{2\pi}{N}}{l}\int_0^{\frac{l}{2v\left(1-\cos\frac{2\pi}{N}\right)}} \frac{1}{1+\frac{v}{l}\left(\cos\frac{2\pi}{N}-1\right)t}dt = \frac{\sin\frac{2\pi}{N}}{1-\cos\frac{2\pi}{N}}\ln 2.$$

17. *Constant Magnitude of Acceleration***

Using α defined in the hint above (with the radial component of acceleration pointing inwards when α is positive), the kinematics equation for circular motion in polar coordinates is

$$-a\sin\alpha = -r\dot{\theta}^2,$$
$$a\cos\alpha = r\ddot{\theta}.$$

Differentiating the first equation with respect to time t,

$$-a\cos\alpha\dot{\alpha} = -2r\dot{\theta}\ddot{\theta}.$$

Dividing this by the second equation,

$$\dot{\alpha} = 2\dot{\theta}.$$

α starts off at zero (as the particle does not require any centripetal acceleration yet) and ends at $\frac{\pi}{2}$ as the acceleration must point purely radially inwards when the particle's angular velocity is maximum. Integrating the above,

$$\frac{\pi}{2} = 2\Delta\theta$$
$$\Delta\theta = \frac{\pi}{4}.$$

Thus, the required distance travelled by the particle is $\frac{\pi r}{4}$.

18. *Searching in Fog****

Due to the impossibility of identifying the direction of the pirate ship's velocity, you have to search the entire 2π radians in polar coordinates in the minimum amount of time — for which the following plan is optimal. Define the origin to be at your initial position. First, you move in an arbitrary direction for a distance l_0 which corresponds to the distance you need to cover if the pirate ship indeed traveled in that direction. After covering this distance, you can guarantee that the pirate ship did not follow this direction. Afterwards, you travel tangentially and radially such that your radial coordinate r follows

$$r = l_0 + vt,$$

where t is the time after you have traveled a distance l_0. If your angular position is θ at time t, you can eliminate the possibility of the angular direction of the pirate ship's velocity being θ. As you cover $\theta = 0$ to $\theta = 2\pi$ with your non-zero tangential velocity, you would have considered all possibilities and are guaranteed to have found the ship. Now, we proceed to calculate the required time for this strategy. Due to the time delay t_0, the initial distance between you and the pirate ship is vt_0. The time taken for you to catch up, if you anticipated its direction of motion correctly is

$$\tau = \frac{vt_0}{u - v},$$

$$l_0 = u\tau = \frac{uvt_0}{u - v}.$$

We have computed the required time for the first process. For the second process, your radial and tangential velocities are related by

$$\dot{r}^2 + r^2\dot{\theta}^2 = u^2.$$

Since $r = l_0 + vt \implies \dot{r} = v$,

$$\dot{\theta} = \frac{\sqrt{u^2 - v^2}}{l_0 + vt}$$

$$\int_0^{2\pi} v d\theta = \int_0^{\tau'} \frac{\sqrt{u^2 - v^2}}{\frac{l_0}{v} + t} dt$$

$$2\pi v = \sqrt{u^2 - v^2} \ln \left| \frac{l_0 + v\tau'}{l_0} \right|$$

$$\tau' = \frac{l_0}{v}\left(e^{\frac{2\pi v}{\sqrt{u^2 - v^2}}} - 1\right) = \frac{ut_0}{u - v}\left(e^{\frac{2\pi v}{\sqrt{u^2 - v^2}}} - 1\right).$$

The total required time is

$$t_{total} = \tau + \tau' = \frac{ut_0}{u - v} e^{\frac{2\pi v}{\sqrt{u^2 - v^2}}} - t_0.$$

19. *Two Velocities**

The instantaneous center of rotation must be the point of intersection of the lines passing through the two points that are perpendicular to their velocities. In this case, the instantaneous center of rotation must be at infinity, implying that the angular velocity of the rigid body is zero else v will be infinite. Hence, all points on the body travel at the same velocity $\boldsymbol{v}$.

20. *Two Plates**

Let v_{CM} denote the velocity of the center of mass, positive rightwards. Then,

$$v_{CM} - \omega R = v_1,$$

$$v_{CM} + \omega R = v_2,$$

$$v_{CM} = \frac{v_1 + v_2}{2},$$

$$\omega = \frac{v_2 - v_1}{2R},$$

where ω is positive clockwise.

21. *Pulling a Square**

Define the x and y-axes to be positive rightwards and upwards. Consider the two rods above. Let the angular velocities of the left and right rods be ω_1 and ω_2 respectively, positive anti-clockwise. Then, we can apply Eq. (3.25) twice to relate v_1 to v_2 in terms of the angular velocities.

$$\begin{pmatrix} v_2 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \omega_1 \end{pmatrix} \times \begin{pmatrix} \frac{\sqrt{2}}{2}l \\ \frac{\sqrt{2}}{2}l \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \omega_2 \end{pmatrix} \times \begin{pmatrix} \frac{\sqrt{2}}{2}l \\ -\frac{\sqrt{2}}{2}l \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} v_1 - \frac{\sqrt{2}}{2}l\omega_1 + \frac{\sqrt{2}}{2}l\omega_2 \\ \frac{\sqrt{2}}{2}l\omega_1 + \frac{\sqrt{2}}{2}l\omega_2 \\ 0 \end{pmatrix}.$$

Comparing the first and second entries,

$$\omega_1 = -\omega_2$$

$$v_1 - \sqrt{2}\omega_1 l = v_2$$

$$\omega_1 l = \frac{v_1 - v_2}{\sqrt{2}}.$$

The velocity of the top pin is then

$$\begin{pmatrix} v_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \omega_1 \end{pmatrix} \times \begin{pmatrix} \frac{\sqrt{2}}{2}l \\ \frac{\sqrt{2}}{2}l \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{v_1+v_2}{2} \\ \frac{v_1-v_2}{2} \\ 0 \end{pmatrix}.$$

A slick solution would be to consider a frame that travels at velocity $\frac{v_1+v_2}{2}$ with respect to the lab frame. Then, the velocities of the left and right pins are $\frac{v_1-v_2}{2}$ and $-\frac{v_1-v_2}{2}$. Due to the symmetry of this set-up, the top pin can only have a vertical velocity, whose value must be $\frac{v_1-v_2}{2}$ upwards to maintain the rigid body condition, in this frame. Therefore its velocity in the original frame is $(\frac{v_1+v_2}{2}, \frac{v_1-v_2}{2}, 0)$.

22. *Moving Wheel**

Using the coordinate system defined in the problem, we find that

$$\tan\theta = \frac{r}{x}.$$

Differentiating with respect to time,

$$\sec^2\theta\dot{\theta} = -\frac{r}{x^2}v.$$

Using the fact that $\tan^2\theta + 1 = \sec^2\theta$,

$$\dot{\theta} = -\frac{r}{x^2\sec^2\theta}v = -\frac{r}{r^2+x^2}v.$$

Another method that does not use differentiation is to observe that $\dot{\theta}$ is the tangential velocity $-\frac{v}{\sin\theta}$ (negative as it tends to reduce θ) divided by the radial distance $\frac{x}{\cos\theta}$ such that $\dot{\theta} = -\frac{v}{x}\sin\theta\cos\theta = -\frac{v}{x}\cdot\frac{rx}{r^2+x^2} = -\frac{r}{r^2+x^2}v$.

23. *H-Shape***

If the particle has not exceeded the right-most point of the H-shape after it has rotated by $\frac{\pi}{4}$ radians, the particle will not be able to escape the structure. This can be easily visualized through Fig. 3.29.

If the particle is still within the structure at this instant, the H-shape will rotate further and block off all paths of exit in the subsequent motion.

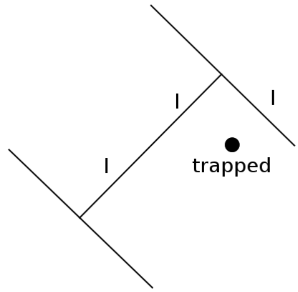


Figure 3.29: Configuration after a $\frac{\pi}{4}$ radians rotation

Furthermore, the particle would have definitely collided with the middle segment of the H-shape after it has rotated by $\frac{\pi}{2}$ radians. Thus, the particle must have traveled at least $\sqrt{2}l$ rightwards, relative to the center of mass of the H-shape by the time it rotates by $\frac{\pi}{4}$ radians. This requires

$$\frac{\pi}{4\omega}(v - v_{CM}) \geq \sqrt{2}l$$

$$v \geq v_{CM} + \frac{4\sqrt{2}\omega l}{\pi}.$$

24. *Inverse-Squared Force**

Using $\ddot{r} = \dot{r}\frac{d\dot{r}}{dr}$ and separating variables,

$$\int \dot{r} d\dot{r} = \int \frac{k}{r^2} dr$$

$$\dot{r} = \pm\sqrt{c - \frac{2k}{r}},$$

where c is a constant.

25. *Bernoulli's Equation***

$z = y^{1-n}$ implies that

$$z' = (1-n)y^{-n}y',$$

where a prime denotes differentiation with respect to x. Dividing the original equation by y^n,

$$y'y^{-n} + f(x)y^{1-n} = g(x)$$

$$\frac{z'}{1-n} + f(x)z = g(x).$$

Multiplying the above by $1 - n$,

$$z' + (1-n)f(x)z = (1-n)g(x).$$

Hence, the appropriate integrating factor for this equation is

$$e^{\int (1-n)f(x)dx}.$$

26. *Bug on Rubber Band****

Let $x(t)$ be the distance between the bug and the right end of the rubber band at time t. At time t, the length of the rubber band is $l + vt$. Therefore, at time $t + dt$,

$$x + dx = x \cdot \frac{l + v(t+dt)}{l + vt} - udt.$$

In the time interval dt, the rubber band is stretched by a distance vdt which causes x to increase by a factor corresponding to the first term on the right-hand side. Furthermore, the bug also crawls a distance udt in this time interval — decreasing x by udt. The exact order of these two events does not matter as the differences are second-order. Simplifying the above equation,

$$\dot{x} - \frac{v}{l+vt}x = -u.$$

Multiplying the above by the appropriate integrating factor $\frac{v}{l+vt}$,

$$\frac{v}{l+vt}\dot{x} - \frac{v^2}{(l+vt)^2}x = \frac{d\left(\frac{v}{l+vt}x\right)}{dt} = -\frac{uv}{l+vt}.$$

Separating variables and integrating,

$$\int_v^{\frac{v}{l+vt}x} d\left(\frac{v}{l+vt}x\right) = -\int_0^t \frac{u}{\frac{l}{v}+t}dt$$

$$\frac{v}{l+vt}x - v = -u\ln\left|\frac{l+vt}{l}\right|$$

$$x = l + vt - \frac{u(l+vt)}{v}\ln\left|\frac{l+vt}{l}\right|.$$

When $x = 0$,

$$t = \frac{l}{v}\left(e^{\frac{v}{u}} - 1\right).$$

Chapter 4

Translational Dynamics

In the next few chapters, we will explore the dynamical laws that describe the behavior of objects under the influence of forces and torques. In this chapter, we will be focusing on forces and Newton's three laws.

4.1 Linear Momentum

The concept of a force should not be discussed in isolation. We shall first introduce the notion of linear momentum. The linear momentum $\boldsymbol{p}$ of a particle is defined as

$$\boldsymbol{p} = m\boldsymbol{v},$$

where m and $\boldsymbol{v}$ are the particle's mass and velocity respectively. The total momentum of a system of particles is simply the sum of the individual contributions due to each particle. This definition of linear momentum arises from an interesting empirical observation — the conservation of momentum of a system under certain conditions. This will be elaborated on in a later chapter but for now, the reader should understand that linear momentum is a useful and convenient quantity that has certain unique properties.

4.2 Newton's Three Laws

4.2.1 *The First Law and Inertial Frames*

Newton's first law states that objects tend to continue in a state of constant velocity unless acted upon by a net external force. The first law actually defines an inertial frame of reference; an inertial frame of reference is one where the first law holds true. An inertial frame of reference is one that does not possess an absolute acceleration (i.e. an accelerometer placed in that frame will measure no acceleration).

This absoluteness of acceleration is one of the salient features of Newtonian physics, as opposed to velocity which is relative. Consider the following set-up which is similar to Newton's renowned thought experiment involving a rotating bucket of water: suppose that you have a glass of water and are traveling on a train. When the train is moving at a constant velocity, the water surface remains flat. However when the train accelerates, the water surface begins to tilt which reflects the magnitude of acceleration. All external observers, accelerating or moving at a constant velocity, will observe the same tilt of the water surface and conclude that you have an acceleration of a certain magnitude which is consistent with all observers. Thus, an absolute acceleration exists from a Newtonian perspective.

Adding on to the previous definition regarding non-accelerating frames, inertial frames of reference are a set of frames that travel at constant velocity with respect to each other. That is, a frame that travels at a constant velocity relative to an inertial frame is also an inertial frame. Newton hypothesized the existence of an absolute space which is truly stationary. He used the frame of distant stars as a reference from which all other inertial frames could be derived.

Ultimately, the first law establishes the context in which Newton's laws are valid. It is definitely not a limiting case of the second law — a common misconception. To illustrate this, there exist frames in which free particles (i.e. free from forces) experience an acceleration. For example, passengers in an accelerating train will observe their external surroundings to accelerate. Clearly, this "violates" Newton's second law but this is perfectly fine. Newton's laws cannot be applied to this reference frame in the first place as it is not an inertial frame.

4.2.2 *The Second Law*

Newton's second law states that the net external force on a particle $\sum \boldsymbol{F}$ is directly proportional to the rate of change of its momentum $\boldsymbol{p}$:

$$\sum \boldsymbol{F} = \frac{d\boldsymbol{p}}{dt}. \tag{4.1}$$

Substituting the expression for the momentum of a particle,

$$\sum \boldsymbol{F} = \frac{d(m\boldsymbol{v})}{dt} = m\frac{d\boldsymbol{v}}{dt} = m\boldsymbol{a}, \tag{4.2}$$

where $\boldsymbol{a}$ is the acceleration of the particle. For a system of particles, an analogous $\sum \boldsymbol{F} = m\boldsymbol{a}$ equation exists but $\boldsymbol{a}$ refers to a different physical quantity now (this is expected as we cannot simply fit the acceleration of a particular particle into the equation). This shall be analyzed after a discussion of the third law.

4.2.3 *The Third Law*

It is common sense that when you punch a wall, the wall "hits" you back as you feel an impact on your knuckles. The third law formulates this explicitly:

$$\boldsymbol{F}_{AB} = -\boldsymbol{F}_{BA}, \tag{4.3}$$

where $\boldsymbol{F}_{AB}$ refers to the force on object A by object B. Thus when object A applies a force on object B, it also experiences an equal and opposite force due to object B. The first two laws do not explicate the motion of the source of the force but the third law weaves it in intricately. However, keep in mind that there are exceptions to the third law such as the magnetic force.[1]

The third law is also known as the weak law of action and reaction — in direct contrast with the strong law which requires the equal and opposite forces to lie on the same line of action. Most forces such as gravity and the normal force in fact obey the strong law of action and reaction.

4.3 Net External Force on a System of Particles

The combination of Newton's second and third laws implies an elegant restatement of the second law for a system of particles. Let us analyze a system of N particles, while using $(\sum \boldsymbol{F})_i$ to denote the net external force on the ith particle and $\boldsymbol{f}_{ij}$ to denote the internal force on the ith particle due to the jth particle. Applying the second law to the ith particle,

$$\left(\sum \boldsymbol{F}\right)_i + \sum_{j,j\neq i} \boldsymbol{f}_{ij} = \frac{d\boldsymbol{p}_i}{dt}.$$

Summing over all particles,

$$\sum_{i=1}^{N} \left(\sum \boldsymbol{F}\right)_i + \sum_{i,j\, i\neq j} \boldsymbol{f}_{ij} = \frac{d}{dt}\left(\sum_{i=1}^{N} \boldsymbol{p}_i\right). \tag{4.4}$$

The first term represents the total net external force acting on the system $\sum \boldsymbol{F}_{ext}$ while the second term evaluates to zero as Newton's third law states

[1] Consider two positive charges traveling along the x and y-axes respectively. The magnetic forces on the particles due to the other particle are directed along different directions (in the y and x directions). However, linear momentum is still conserved if a momentum is associated with the electromagnetic field.

that

$$\boldsymbol{f}_{ij} = -\boldsymbol{f}_{ji}.$$

Equation (4.4) becomes

$$\sum \boldsymbol{F}_{ext} = \frac{d\boldsymbol{P}}{dt} = \sum_{i=1}^{N} m_i \boldsymbol{a}_i. \tag{4.5}$$

That is, the rate of change of the total momentum of a system of particles $\frac{d\boldsymbol{P}}{dt}$ is equal to the net external force on the system $\sum \boldsymbol{F}_{ext}$. Furthermore, if we define a positional attribute of the system, known as its center of mass, whose position vector is given by

$$\boldsymbol{R} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{r}_i}{\sum_{i=1}^{N} m_i}, \tag{4.6}$$

where $\boldsymbol{r}_i$ is the position vector of the ith particle and the total mass of the system is

$$M = \sum_{i=1}^{N} m_i,$$

then,

$$\boldsymbol{P} = \sum_{i=1}^{N} m_i \frac{d\boldsymbol{r}_i}{dt} = M\frac{d\boldsymbol{R}}{dt} = M\boldsymbol{v}_{CM}, \tag{4.7}$$

where $\boldsymbol{v}_{CM} = \frac{d\boldsymbol{R}}{dt}$ is the velocity of the center of mass. Furthermore, Eq. (4.5) yields

$$\sum \boldsymbol{F}_{ext} = M\frac{d^2\boldsymbol{R}}{dt^2} = M\boldsymbol{a}_{CM}, \tag{4.8}$$

where $\boldsymbol{a}_{CM} = \frac{d^2\boldsymbol{R}}{dt^2}$ is the acceleration of the center of mass. We see that a system of particles translationally responds to a net external force as if it were an imaginary mass M located at the center of mass of the system, $\boldsymbol{R}$. Equation (4.8) has profound ramifications for a rigid body as it implies that we no longer need to write down $\boldsymbol{F} = m\boldsymbol{a}$ for each particle. We can simply analyze the translational motion of the center of mass which defines the translational motion of the entire body, as relative distances between particles must be preserved (the rotational motion about the center of mass still needs to be studied in a different manner to determine the orientation of the rigid body).

For a continuous mass distribution, a similar procedure would lead to the conclusion that

$$\boldsymbol{R} = \frac{\int \boldsymbol{r} dm}{\int dm}, \tag{4.9}$$

where $\boldsymbol{r}$ denotes the position vector of the infinitesimal mass element dm under consideration. That is, $\int \boldsymbol{r} dm$ represents integrating the position vector of each infinitesimal mass element $\boldsymbol{r}$, with weights given by its infinitesimal mass dm, over the entire distribution (the integral could be a line, surface or volume integral).

4.3.1 *Center of Mass*

The computation of the center of mass of a system shall be illustrated with a few examples. We first begin with a simple system of two particles m and M, which are located at coordinates (x_1, y_1, z_1) and (x_2, y_2, z_2) respectively. From Eq. (4.6),

$$\boldsymbol{R} = \frac{1}{m+M}\begin{pmatrix} mx_1 \\ my_1 \\ mz_1 \end{pmatrix} + \frac{1}{m+M}\begin{pmatrix} Mx_2 \\ My_2 \\ Mz_2 \end{pmatrix} = \frac{1}{m+M}\begin{pmatrix} mx_1 + Mx_2 \\ my_1 + My_2 \\ mz_1 + Mz_2 \end{pmatrix}.$$

Next, as a consequence of the linearity of $\boldsymbol{R}$, a trick in calculating the center of mass of a system of discrete particles or for a system with multiple continuous mass distributions is to calculate the center of mass of a few selected particles or distributions and replace them with a point mass, equal to the sum of their masses, at that exact position. Then, you can proceed to calculate the center of mass of the whole system with the point mass replacing some of the original particles or distributions. The proof of this shall be left as an exercise to the reader.

Now, let us apply this sleight-of-hand to another simple example.

Suppose we wish to calculate $\boldsymbol{R}$ for the system of two uniform spheres depicted in Fig. 4.1. No amount of ingenuity in the selection of coordinates will lead to a simple integration with continuous limits. However, we can use symmetrical arguments[2] to argue that the center of mass of each individual

[2] The center of mass must lie along lines/planes of symmetry of a uniform mass distribution as a direct corollary of its definition. For example, in the two-dimensional case, define the origin along a line of symmetry (we define the y-axis to be parallel to this line) and consider the direction perpendicular to this line (x direction). For every particle at coordinates (x, y), there will be a corresponding particle at coordinates $(-x, y)$ due to symmetry. The weighted sum of the x-coordinates of the distribution must consequently

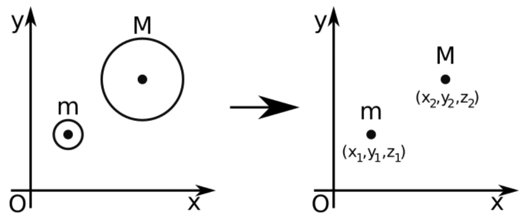


Figure 4.1: Center of mass of two spheres

sphere must be at its center and then reduce the system to two point masses, m and M, at the corresponding centers. Then we can directly apply the result of the previous problem.

The above method can also be applied to problem involving "missing masses." Firstly, wisely introduce imaginary masses to the system. Subsequently, replace the combined system, comprising the original system and the imaginary masses, with the total mass of the combined system at its center of mass and a negative mass, commensurate with the imaginary masses that have been added, at the center of mass of the imaginary system. Then, the center of mass of this new set-up is equivalent to that of the original set-up.

Problem: Referring to Fig. 4.2, a uniform circle of radius R has a circular hole of radius $\frac{R}{2}$. If the uniform surface mass density is σ, determine the center of mass of the system.

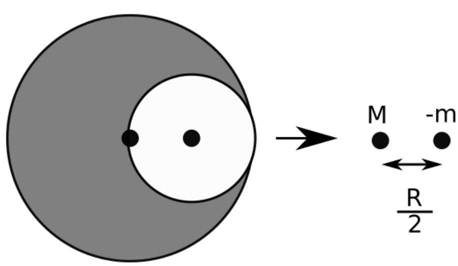


Figure 4.2: Circle with a hole

To determine the center of mass of this object, we can first "fill up" the hole to produce a complete circle which can be reduced to a mass $M = \sigma\pi R^2$

be zero and hence, the center of mass must lie along this line of symmetry. As seen from this example, the intersection of lines and planes of symmetry can be used to determine the center of mass.

at its center. To counteract the effect of this additional mass, contributions due to the mass in the hole must be subtracted. This is equivalent to adding "negative mass" in the region of the hole which is equivalent to a mass $-m = -\frac{1}{4}\sigma\pi R^2$ at the center of the hole. If we define the origin to be at the center of the complete circle, x_{CM} of the circle with a hole is

$$x_{CM} = \frac{-m \cdot \frac{R}{2}}{M - m} = -\frac{R}{6}.$$

Center of Mass of a Continuous Distribution

For continuous mass distributions, integration is generally required to determine the center of mass. The main mathematical difficulty in this process concerns choosing a convenient coordinate system and adopting the correct limits of integration that physically correspond to the mass distribution. Chapter 2 analyzes these aspects in greater detail. Hopefully, the following two examples can provide some form of intuition.

Problem: Determine the center of mass of a right-angled triangle of a uniform surface mass density σ and with an angle of inclination θ in the xy-plane. Let the length of its base be l.

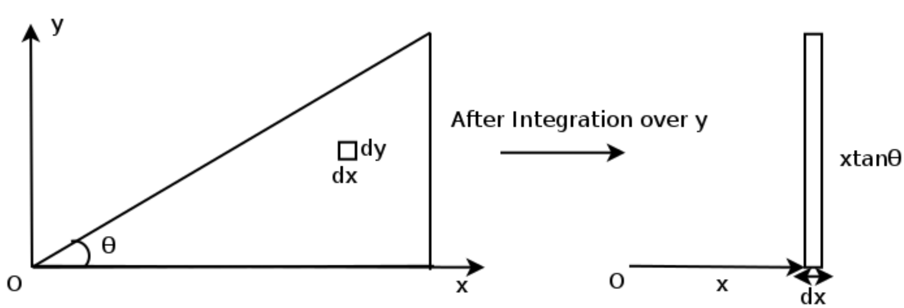


Figure 4.3: Center of mass of a triangle

To solve this problem, we consider infinitesimal rectangular elements in Cartesian coordinates with sides dx and dy. Then, $dm = \sigma dxdy$. Applying Eq. (4.9), we need to perform the double integral

$$\boldsymbol{R} = \frac{1}{M} \iint_S \begin{pmatrix} x \\ y \end{pmatrix} \sigma dxdy$$

over the surface S, which is the triangle in this case. M, as always, refers to the total mass of the triangle. This is a case where the limits of integration depend on the order of integration. We can choose to integrate over y before x. Diagrammatically, this corresponds to first integrating over a vertical strip of mass at a particular x-coordinate x (right of Fig. 4.3) before integrating over all strips. The limits for y for a vertical strip at horizontal coordinate x are evidently 0 and $x \tan\theta$ while the limits of x are 0 and l.

$$\boldsymbol{R} = \frac{1}{M} \int_0^l \int_0^{x\tan\theta} \begin{pmatrix} x \\ y \end{pmatrix} \sigma dy dx.$$

Performing the integrals for the two coordinates separately,

$$x_{CM} = \frac{\int_0^l \int_0^{x\tan\theta} x\sigma dy dx}{M} = \frac{\int_0^l x^2 \tan\theta \sigma dx}{M} = \frac{l^3 \tan\theta \sigma}{3M} = \frac{2l}{3},$$
$$y_{CM} = \frac{\int_0^l \int_0^{x\tan\theta} y dy dx}{M} = \frac{\int_0^l x^2 \tan^2\theta \sigma dx}{2M} = \frac{l^3 \tan^2\theta \sigma}{6M} = \frac{l\tan\theta}{3},$$

where the last equalities are obtained from using $M = \frac{l^2 \tan\theta\sigma}{2}$. Note that we could have stopped after calculating x_{CM} and argued that the y-coordinate of the center of mass of the triangle should be one-third of the y-coordinate of the right tip by symmetry (as the result for x_{CM} is independent of θ). Observe that the center of mass, in this case, is the intersection of the three medians which divides each median into two segments in the ratio 2:1 (the shorter segment is closer to the base). This is in fact true for a general triangle and is a well-known geometrical result.

Problem: Determine the center of mass of a semi-circle with a uniform surface mass density σ and radius R.

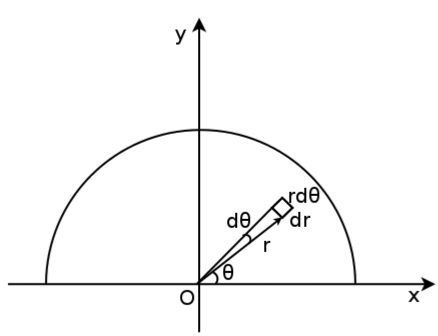


Figure 4.4: Semi-circle

It is convenient to adopt polar coordinates. Then, the infinitesimal mass element at polar coordinates (r, θ) is a rectangle of sides dr and $rd\theta$ ($dm =$

$\sigma r d\theta dr$). Its position vector $\boldsymbol{r}$ in Cartesian coordinates is $(r\cos\theta, r\sin\theta)$. This step of expressing the position coordinate in terms of Cartesian coordinates is pivotal, because an integration over the perpetually changing basis vectors in polar coordinates is extremely cumbersome. Following from this,

$$\boldsymbol{R} = \frac{1}{M}\int_0^R \int_0^\pi \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix} \sigma r d\theta dr.$$

We shall just evaluate y_{CM} as x_{CM} is zero due to symmetry.

$$y_{CM} = \frac{1}{M}\int_0^R \int_0^\pi r^2 \sin\theta \sigma d\theta dr = \frac{1}{M}\int_0^R 2r^2\sigma dr = \frac{2\sigma R^3}{3M} = \frac{4R}{3\pi}.$$

Rotations

Actually, for rotationally symmetric objects, such as the semi-circle in the previous section, we can visualize the effects of rotating the object by a small angle ϕ (so that some mass is transferred from one end to another end) to determine the center of mass. This change is best illustrated via the change in potential energy in a uniform gravitational field (though it is solely due to the changes in the coordinates of the masses and has nothing to do with the dynamics of the system).

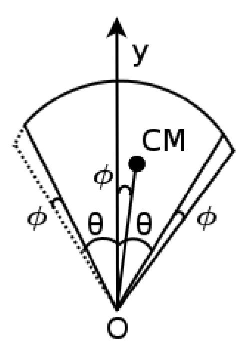


Figure 4.5: Rotated sector

Let's say we wish to determine y_{CM} of a uniform sector of angle 2θ and radius R that is initially placed symmetrically about the y-axis and centered at the origin. Suppose that we rotate the sector by a small clockwise angle ϕ. A thin isosceles triangle of equal sides R and apex angle ϕ is effectively displaced from the left to the right (Fig. 4.5). Since the centre of mass of an isosceles triangle is known to be at two-thirds of its height to the lone base with a different length,[3] the change in the vertical height of center of mass of the small triangle is

$$\Delta y = \frac{2}{3}R\left[\cos\left(\theta + \frac{\phi}{2}\right) - \cos\left(\theta - \frac{\phi}{2}\right)\right] = \frac{2}{3}R \cdot -2\sin\theta\sin\frac{\phi}{2} \approx -\frac{2}{3}R\sin\theta\phi,$$

as $\sin x \approx x$ for small x. Since the area of the triangle is $\frac{1}{2}R^2\sin\phi \approx \frac{1}{2}R^2\phi$ by the sine rule, the change in the potential energy of the entire sector due to this rotation is

$$\Delta U = \sigma\frac{1}{2}R^2\phi \cdot g \cdot -\frac{2}{3}R\sin\theta\phi = -\frac{1}{3}\sigma R^3 g\sin\theta\phi^2,$$

where $\sigma\frac{1}{2}R^2\phi$ is the mass of the triangle. We can evaluate ΔU in another way[4] by considering the change in the vertical coordinate of the center of mass of the entire sector.

$$\Delta U = -Mgy_{CM}(1 - \cos\phi),$$

where M is the total mass of the sector. Substituting $M = \theta\sigma R^2$ and applying the small angle approximation $\cos x \approx 1 - \frac{x^2}{2}$,

$$\Delta U = -\frac{1}{2}\theta\sigma R^2 g y_{CM}\phi^2.$$

Equating the two expressions for ΔU,

$$y_{CM} = \frac{2R\sin\theta}{3\theta}.$$

Substituting $\theta = \frac{\pi}{2}$, we retrieve $y_{CM} = \frac{4R}{3\pi}$ for a semi-circle.

Scaling Arguments

If we scale all length dimensions of an object by a factor k, the distances between two corresponding points will also be scaled by a factor of k. This

[3]An isosceles triangle can be seen as the composition of two right-angled triangles whose center of masses have been proven to be located at coordinates corresponding to two-thirds of the lengths of the non-hypotenuse sides.

[4]This is because the gravitational potential energy of an extended body in a uniform gravitational field is equivalent to that of a point mass, commensurate with the total mass of the extended body, placed at its center of mass. To prove this, the total gravitational potential energy of an extended body is $\int gr_y dm$ where r_y is the y-coordinate of the position vector of an infinitesimal mass element dm on the extended body and where the integral is performed over the entire body. Since g is uniform, $\int gr_y dm = g\int r_y dm = Mgy_{CM}$ where M is the total mass of the extended body. The last equality comes from the definition of the center of mass.

fact can be used to determine the center of mass of appropriate objects without any integration. Consider the example of a right-angled triangle again.

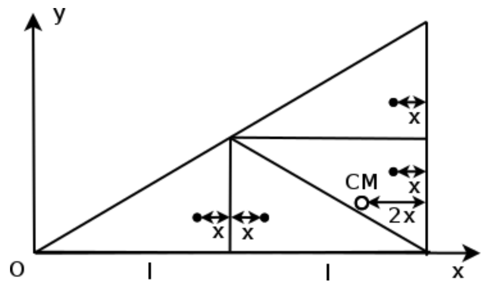


Figure 4.6: Triangle of base length $2l$

Let the horizontal position of the center of mass of a right-angled triangle with angle of inclination θ and base length l be located at a horizontal distance x away from its height. Then, that of a right-angled triangle with base length $2l$ will be $2x$ (depicted by the hollow white circle labeled "CM" in Fig. 4.6). Furthermore, this larger triangle is composed of 4 original triangles as shown in the diagram. Thus, we can replace each of the smaller triangles with a point mass at its center of mass which is illustrated by a black dot (the exact vertical coordinate does not matter for now as we only wish to determine the horizontal coordinate) and evaluate X_{CM} for the larger triangle. Using these different methods to calculate X_{CM}, we obtain an equation in x.

$$X_{CM} = \frac{l-x}{4} + \frac{l+x}{4} + \frac{2l-x}{4} + \frac{2l-x}{4} = 2l - 2x.$$

The first expression for X_{CM} is obtained from considering the weighted contributions due to the four smaller triangles while the second expression is obtained from scaling arguments. Then,

$$x = \frac{l}{3}.$$

The x-coordinate of the center of mass of the triangle with base length l is then

$$x_{CM} = l - x = \frac{2l}{3}.$$

A similar argument can be made for the y-coordinate of the center of mass.

4.4 Equations of Motion in Different Coordinates

The $\sum \boldsymbol{F} = m\boldsymbol{a}$ equations are inherently vector equations. To procure meaning from these equations, we often need to express them in terms of scalars and then solve the resulting differential equations. This, in turn, depends on the coordinate system utilized.

4.4.1 *Cartesian Coordinate System*

The net force and position vector in a Cartesian system can be expressed as

$$\sum \boldsymbol{F} = F_x\hat{\boldsymbol{\imath}} + F_y\hat{\boldsymbol{\jmath}} + F_z\hat{\boldsymbol{k}},$$
$$\boldsymbol{r} = x\hat{\boldsymbol{\imath}} + y\hat{\boldsymbol{\jmath}} + z\hat{\boldsymbol{k}}.$$

The acceleration is defined as

$$\boldsymbol{a} = \frac{d^2\boldsymbol{r}}{dt^2}.$$

Note that the rate of a change of a vector $\boldsymbol{A}$ depends on both the change in magnitude and direction of the vector($\frac{dA\hat{\boldsymbol{A}}}{dt} = \frac{dA}{dt}\hat{\boldsymbol{A}} + A\frac{d\hat{\boldsymbol{A}}}{dt}$). Since the basis vectors $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$ are fixed,

$$\boldsymbol{a} = \frac{d^2x}{dt^2}\hat{\boldsymbol{\imath}} + \frac{d^2y}{dt^2}\hat{\boldsymbol{\jmath}} + \frac{d^2z}{dt^2}\hat{\boldsymbol{k}}.$$

Considering the corresponding components of $\sum \boldsymbol{F} = m\boldsymbol{a}$ yields

$$F_x = \frac{d^2x}{dt^2},$$
$$F_y = \frac{d^2y}{dt^2},$$
$$F_z = \frac{d^2z}{dt^2}.$$

4.4.2 *Polar Coordinate System*

In a two-dimensional polar coordinate system, the basis vectors are $\hat{\boldsymbol{r}}$, the unit vector pointing from the origin to the position of the point of concern, and $\hat{\boldsymbol{\theta}}$, a unit vector tangential to $\hat{\boldsymbol{r}}$. The net force and position vector in a

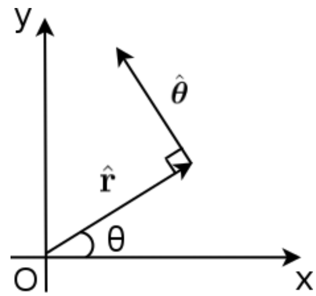


Figure 4.7: Polar coordinates

polar coordinate system are

$$\sum \boldsymbol{F} = F_r \hat{\boldsymbol{r}} + F_\theta \hat{\boldsymbol{\theta}},$$
$$\boldsymbol{r} = r\hat{\boldsymbol{r}}.$$

The acceleration is

$$\boldsymbol{a} = \frac{d^2\boldsymbol{r}}{dt^2}.$$

Note that basis vectors are now variable and their change must be considered when the derivative of the position vector is evaluated. The acceleration, expressed in terms of the instantaneous basis vectors, was derived in Chapter 3 to be

$$\boldsymbol{a} = (\ddot{r} - r\dot{\theta}^2)\hat{\boldsymbol{r}} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{\theta}}.$$

Comparing the corresponding components of $\sum \boldsymbol{F} = m\boldsymbol{a}$ yields

$$F_r = m(\ddot{r} - r\dot{\theta}^2), \tag{4.10}$$

$$F_\theta = m(r\ddot{\theta} + 2\dot{r}\dot{\theta}). \tag{4.11}$$

4.5 Typical Forces in Mechanics

Now that we can determine the evolution of a system given the forces on it, we shall focus on evaluating the forces. There are four fundamental interactions in nature, namely: gravitational, electromagnetic, strong nuclear and weak nuclear. All mechanical forces — which require direct contact to be delivered — are in fact electromagnetic. These include the normal force, friction and the spring force. When you push on a door, the electrons in your fingertips are compressed against those on the door — resulting in repulsion and a normal force. As another example, the tension in a string arises from the mutual attraction between atoms of the string which tends to prevent the string from disintegrating when it is stretched. The only fundamental force in this section is in fact the gravitational force.

4.5.1 *Normal Force*

The normal force, as its name implies, is the force that acts perpendicular to the surface between two objects. When we are standing on the ground, the normal force on us due to the ground opposes the gravitational force and prevents us from accelerating towards the core of the Earth. Note that the normal force is a contact force and it acts at the point of contact between objects. This is particularly important when we learn about torques later. Finally, note that an object's apparent weight is the normal force exerted by the object on an imaginary weighing scale or vice-versa.

4.5.2 *Friction*

Friction is a force that resists relative motion between two surfaces. It arises from various factors such as surface adhesion, surface deformation and irregular surfaces. We should differentiate between static friction and kinetic friction. Static friction is present when there is no relative movement between two surfaces and is governed by the equation: $|f_s| \leq \mu_s N$ where $|f_s|$ is the magnitude of static friction, μ_s is the coefficient of static friction and N is the normal force between the surfaces. The direction of static friction is oriented such that it opposes impending motion. If a force exceeds the upper limit, the object will begin to move. Kinetic friction, on the other hand, opposes the motion of surfaces that are already relatively moving. It is constant and has a magnitude of $f_v = \mu_v N$. In most cases, $0 < \mu < 1$ and $\mu_v < \mu_s$. However, there are also exceptions such as silicone rubber surfaces which often have coefficients of friction substantially larger than 1. Lastly, it is pivotal to understand that since kinetic friction always acts in a direction opposite to that of motion, it is a non-conservative[5] and dissipative force. Energy is often converted to heat and sound which are transmitted to the surroundings.

4.5.3 *Spring Force*

When a spring, or elastic object in general, is compressed or elongated, there is a tendency for it to rebound to its natural state. This restoring force is proportional to the magnitude of x, the extension/compression of a spring relative to its relaxed length ($F = kx$) by Hooke's law, and is

[5]This concept shall be elaborated on further in Chapter 6.

opposite in direction such that the spring tends to return to its relaxed length. Let us digress for a while and discuss the effective spring constants for massless springs with identical relaxed lengths, arranged in series and parallel configurations. Consider a system of two springs. Let the displacements of the object, the first and second spring from their respective equilibrium positions be x, x_1 and x_2, respectively.

Parallel Configuration

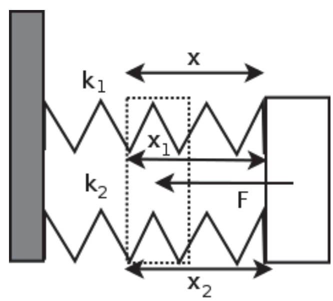


Figure 4.8: Springs connected in parallel

It is evident from Fig. 4.8 that the displacements of the springs and the object must be equal in a parallel configuration.

$$x = x_1 = x_2.$$

Furthermore, the total force on the object is

$$F = k_1x_1 + k_2x_2 = (k_1 + k_2)x.$$

To determine the effective spring constant, we need to find a k_{eff} such that

$$F = k_{\mathit{eff}}\, x$$
$$\implies k_{\mathit{eff}} = k_1 + k_2.$$

Applying this formula repeatedly between pairs of springs, the effective spring constant for a parallel configuration of n springs in general is

$$k_{\mathit{eff}} = \sum_{i=1}^{n} k_i.$$

Series Configuration

For springs connected in series, the force due to each of them must be the same. If not, there will be a net force on the massless springs which will cause them to undergo infinite acceleration in Fig. 4.9.

$$F = k_{\mathit{eff}} x = k_1 x_1 = k_2 x_2.$$

Furthermore, the sum of the displacements of the springs must be that of the object.

$$x_1 + x_2 = x.$$

Solving,

$$\begin{aligned} k_1 k_2 x_1 &= k_2 k_{\mathit{eff}} x \\ k_1 k_2 x_2 &= k_1 k_{\mathit{eff}} x \\ k_1 k_2 (x_1 + x_2) &= (k_1 + k_2) k_{\mathit{eff}} x \\ \implies k_{\mathit{eff}} &= \frac{k_1 k_2}{k_1 + k_2} \quad \text{or} \quad \frac{1}{k_{\mathit{eff}}} = \frac{1}{k_1} + \frac{1}{k_2}. \end{aligned}$$

Repeatedly applying this formula, the effective spring constant for a series configuration in general is

$$\frac{1}{k_{\mathit{eff}}} = \sum_{i=1}^{n} \frac{1}{k_i}.$$

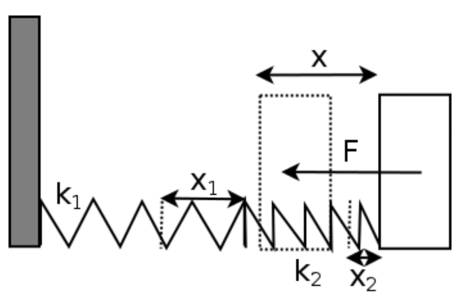


Figure 4.9: Springs connected in series

4.5.4 *Tension*

Tension is the pulling force exerted on one part of a string by its adjacent parts and at the ends of the string. You can imagine the string or rope as an elongated spring and the tension as the restoring force. However, this analogy fails when the string experiences a compressive force as it will become slack (it does not act like a compressed spring). The tension at a segment of a string is always directed along the instantaneous gradient there as it is unable to withstand any forces perpendicular to it. Note that in most cases, strings and ropes are assumed to be massless and thus there must not be any net force acting on them. This condition, in fact, ensures that the tension in a frictionless, massless string of an arbitrary shape is uniform as we shall see later in this chapter. Another proof will also be provided in Chapter 7.

4.5.5 *Gravitational Force*

The gravitational force is a force of attraction between all particles with mass. In the more general Newtonian form, the magnitude of the gravitational force between two point masses is proportional to the product of their masses and inversely proportional to the square of the distance between them. For the rest of the chapter, the only massive object that we will be considering will be the Earth, and differences in height are negligible when compared to the radius of the Earth. Thus, we can assume that there is a gravitational force or weight of mg acting on a object with mass m, where g is the gravitational field strength constant. For an object with non-negligible volume, the center of gravity of the object is defined as the point where the entire force of gravity seems to act on. It is equal to the center of mass when the gravitational field strength is uniform over the entire object (try to prove this).

4.6 Types of Problems

4.6.1 *Free-Body Diagrams*

Before considering the common types of problems, a paramount problem-solving tool is a free-body diagram drawn with the following procedure:

(1) Isolate the system in question.
(2) Identify all external forces acting on the system and draw them as vectors at appropriate points on the system (e.g. the gravitational force should pass through the center of gravity). External forces refer to forces on the system due to entities outside the system.

Then, the general procedure to solving a mechanics problem is as follows.

(1) Consider various systems and draw their free-body diagrams. A useful tip in choosing a system to consider would be to examine the forces you need to solve. If they do not include internal forces (friction and normal force between surfaces and occasionally, tension), you should consider paired objects as a whole system. If you are required to solve for friction or the normal force, you should segregate the objects.
(2) Write the $\sum \boldsymbol{F} = m\boldsymbol{a}$ equations for all systems that you have considered.

(3) If necessary, identify certain relationships between variables (constraints that need to be obeyed) to formulate additional equations so that you have enough equations to solve for the required variables. In some cases, the constraints[6] can also be weaved into the coordinates that define the state of each system.
(4) Solve!

Let us apply this procedure to some problems.

4.6.2 *No Constraints*

Problem: A man of mass m is in a lift that is undergoing an acceleration a upwards. Find the normal force exerted on the man by the floor of the lift (apparent weight).

Let us draw a free-body diagram[7] of our incredibly enthusiastic man taking an elevator. We isolate the man as we would like to find the normal force.

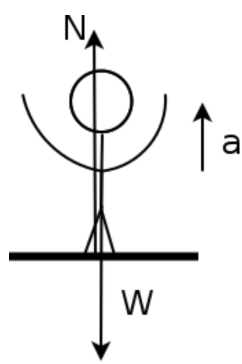


Figure 4.10: A man in an elevator

From Newton's second law, we obtain:

$$N - mg = ma \implies N = m(a+g).$$

Problem: A force is applied on the left end of an arrangement of n identical blocks of mass m that are connected by strings. The coefficient of kinetic

[6]Such constraints are known as holonomic constraints and only depend on the positions of objects. This concept will be further explored in Chapter 11.

[7]The normal force slightly deviates from the accurate position for the sake of clarity. We also draw a on the side to remind ourselves that the man is accelerating upwards. This is not really necessary.

friction between the blocks and the ground is μ. All strings remain taut in the process of the blocks' motion. Find the tension T_i in the string between the ith and $(i+1)$th blocks as the blocks accelerate.

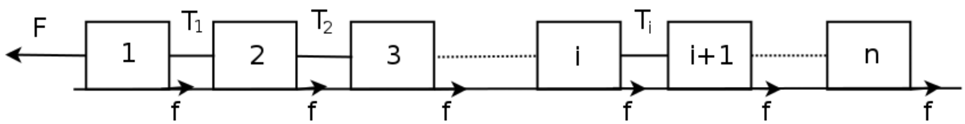


Figure 4.11: Connected blocks

Let f be the frictional force on one block. We first solve for the acceleration of each block by considering all the blocks as a whole system. The acceleration of each block must be the same as the string between them remains taut.

$$F - nf = nma.$$

Next, it is convenient to consider one of our systems as the $(i+1)$th block to the nth block as T_i is the only tension that acts on this system. Choosing a system to isolate requires some intuition which becomes keener with practice. We draw the free-body diagram as follows.

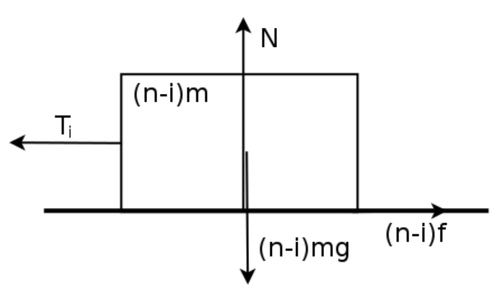


Figure 4.12: Last (n-i) blocks

$$T_i - (n-i)f = (n-i)ma.$$

Solving,

$$T_i = (n-i)(f+ma) = \frac{(n-i)F}{n}.$$

Notice that we did not even need to evaluate f at all! This implies that if we conducted this experiment on the moon for example — such that g and thus f varies — we will obtain the exact same result. Finally, an equally convenient system of blocks that deserves mention is the first i blocks which experience force F, tension $-T_i$, friction $-if$ and weight img.

$$F - T_i - if = ima$$
$$T_i = F - i(f + ma) = F - \frac{iF}{n} = \frac{(n-i)F}{n}.$$

Slinkies

A simple yet fascinating system is an amusing toy known as a slinky, which is basically a massive spring. A major complication in this system arises from the fact that the spring force is not necessarily uniform throughout the slinky as individual segments are no longer massless. The spring is not stretched uniformly such that its density also varies in space, even though its original density may be uniform. In light of the varying spring force, the standard rigorous analysis of such systems involves considering infinitesimal segments of the slinky which are basically smaller massive springs, but we shall see that we can avoid this via some sleights-of-hand.

Problem: A slinky of mass m, spring constant k and relaxed length l — which is the length when tensions at both ends are zero (e.g. placed on a horizontal table) — is hung from a ceiling. Determine the total extension of the slinky and its center of mass after it has attained equilibrium.

Divide the original slinky, before it is stretched, into myriad equal infinitesimal pieces of length dx each. Each piece has the same spring constant k' (which we shall determine later). Without any tedious calculations, we can actually determine the extension of the slinky. The crucial observation is that even though the spring force at equilibrium varies linearly[8] from 0 at the bottom of the slinky to mg at the ceiling (in order to keep all pieces beneath a given piece at equilibrium) such that each piece stretches by a different amount, the extension of each infinitesimal section of the slinky is directly proportional to the tension at its ends. Then, we can exploit the linear variation of tension to conclude that the extension of the slinky is equal to that of a spring with the same spring constant k and uniform tension equal to the average tension $\frac{mg}{2}$. The latter is equivalent to a block of mass $\frac{mg}{2}$ hung onto a massless spring. Therefore, the extension of the slinky is

[8]By linearly, we mean the number accorded to the segment, that we have identified from the original spring, from the bottom and not the length from the bottom.

simply

$$\Delta l = \frac{mg}{2k}.$$

Now, we reach a quandary as we are unable to apply this trick to the second problem (the whole point is to find the center of mass of a spring with a non-uniform mass density). Intuitively, since the tension near the top of the slinky is larger than that at the bottom, the top sections are stretched more than the bottom — causing the density of the slinky to decrease with height and the center of mass to lie below the geometric center. We then proceed with the more rigorous approach. Set the origin at the ceiling and define the x-direction to be positive downwards.

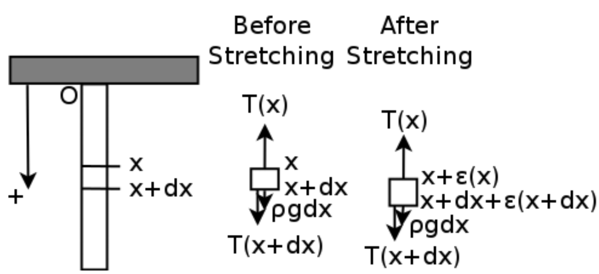


Figure 4.13: Infinitesimal section of slinky

Mentally switch off gravity first, such that the length of the slinky spans points $x = 0$ to $x = l$. Isolate an infinitesimal section between coordinates x and $x + dx$. Since an aggregated spring of spring constant k and length l can be seen as an array of N springs of length $\frac{l}{N}$ in series, the spring constant k_N of each of these smaller springs is

$$k_N = Nk$$

by the series addition formula for springs. In other words, the length l' of a component spring multiplied by its spring constant k' must be equal to kl.

$$k'l' = kl.$$

Applying this to the infinitesimal segment of length dx, its spring constant is

$$k' = \frac{kl}{dx}.$$

Do not worry about the infinitesimal term in the denominator for now. Now, switch on gravity and allow the system to equilibrate. Eventually this infinitesimal section stabilizes between some coordinates $x + \varepsilon(x)$ and $x + dx + \varepsilon(x + dx)$ where $\varepsilon(x)$ represents the increase in the x-coordinate of a point that was originally at coordinate x (Fig. 4.13). The extension of this segment is evidently $\varepsilon(x + dx) - \varepsilon(x)$. The tension as a function of x-coordinate x (remember that x describes the original spring) becomes

$$T = mg\left(1 - \frac{x}{l}\right)$$

to support the weight of the portion below it (those with larger x-coordinates). The tension on the ends of this segment induces its extension according to Hooke's law.

$$T = k'[\varepsilon(x + dx) - \varepsilon(x)] = kl\frac{\varepsilon(x + dx) - \varepsilon(x)}{dx} = kl\frac{d\varepsilon}{dx}.$$

Substituting the expression for T, separating variables and integrating,

$$\int_0^{\varepsilon(x)} d\varepsilon = \int_0^x \frac{mg}{kl}\left(1 - \frac{x}{l}\right) dx,$$

where $\varepsilon(0) = 0$ as the particle at the ceiling is fixed. Simplifying, the total extension of the segment between the ceiling and x-coordinate x is

$$\varepsilon(x) = \frac{mg}{kl}\left(x - \frac{x^2}{2l}\right).$$

One can substitute $x = l$ to check that the total extension of the slinky is indeed $\varepsilon(l) = \frac{mg}{2k}$. Moving on, the center of mass of each infinitesimal segment is basically the coordinates of either of its ends (the difference is second-order and negligible). The final x-coordinate of the top end of the segment which was originally between coordinates x and $x + dx$ is

$$x' = x + \varepsilon(x) = \left(\frac{mg}{kl} + 1\right)x - \frac{mgx^2}{2kl^2}.$$

Note that despite being stretched, this segment still has mass ρdx where $\rho = \frac{m}{l}$ is the original density of the slinky as its boundaries remain the same. To determine the center of mass of the slinky, we simply have to integrate

$$\int_0^l \rho x' dx.$$

Be wary of mistaking dx' for dx as the mass of the section is expressed in terms of ρdx and not its new density multiplied by dx' (you can do so but this step is extraneous).

$$\int_0^l \rho x' dx = \rho \int_0^l \left[\left(\frac{mg}{kl} + 1 \right) x - \frac{mgx^2}{2kl^2} \right] dx = \frac{ml}{2} + \frac{m^2 g}{3k}.$$

Dividing the above by m, we obtain the x-coordinate of the center of mass as

$$x_{CM} = \frac{l}{2} + \frac{mg}{3k},$$

which is $\frac{mg}{3k}$ below the center of the unstretched slinky. Now that we have completed a rigorous discourse on this system, let us present an alternative, elegant approach that leverages on scaling arguments.

The possible parameters of the slinky are m, g, k and l. Based on dimensional analysis, a possible expression for x_{CM} is

$$x_{CM} = \alpha l + \beta \frac{mg}{k},$$

where α and β are some dimensionless constants. Note that even though $\frac{mg}{kl}$ is a dimensionless variable, we do not include arbitrary functions of $\frac{mg}{kl}$ in our guess as the equations describing the system (e.g. Hooke's law and tension as a function of x), are linear and l and $\frac{mg}{k}$ should be independent (we can have a slinky of length $l = 0$ or $m = 0$, for instance, and the same equation should apply. If l and $\frac{mg}{k}$ are coupled, these limiting cases would yield uneventful results.) Hence, it is wise to guess a linear solution in terms of l and $\frac{mg}{k}$. With that said, we can determine α by taking $k \to \infty$ as the slinky becomes so rigid that it hardly budges or deforms. Then,

$$x_{CM} = \frac{l}{2} \implies \alpha = \frac{1}{2},$$

$$\Delta x = x_{CM} - \frac{l}{2} = \beta \frac{mg}{k},$$

where we define a new variable Δx that denotes the excess portion of x_{CM} beyond the x-coordinate of the center of mass of the relaxed slinky. Notice that

$$\Delta x \propto \frac{m}{k}.$$

Therefore, if we hang a slinky of the same linear mass density and length $2l$, the deviation in the x-coordinate of the center of mass from its relaxed state should be $4\Delta x$ as its mass is $2m$ while its spring constant is $\frac{k}{2}$ (2 identical springs in series).

To identify an alternate expression for $4\Delta x$, cut the slinky into two parts of equal masses, which are not of equal equilibrium lengths as the sections

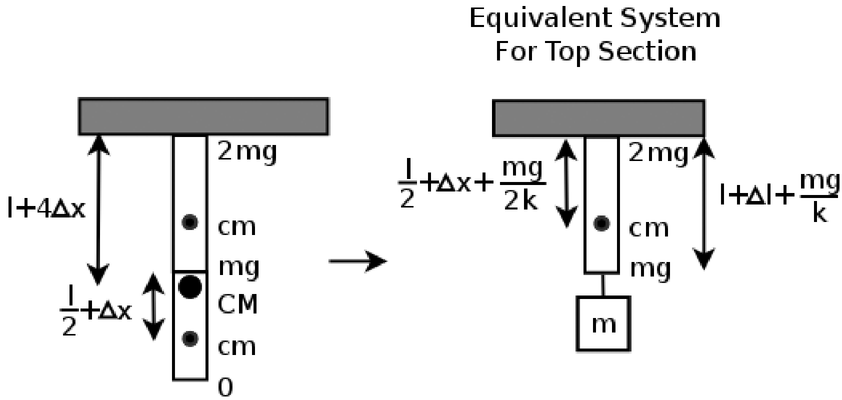


Figure 4.14: Scaling arguments for slinky of length $2l$

near the top are stretched more. The bottom section is a slinky with mass m and original length l with tension mg and 0 at its ends — implying that its center of mass lies a distance $\frac{l}{2}+\Delta x$ from its top end by proposition. The top section is slightly more complex. It is a slinky with mass m and original length l but with tension $2mg$ and mg at its ends as shown in Fig. 4.14. It is equivalent to an original slinky attached with a mass m at its bottom end. The additional weight mg causes the slinky to stretch uniformly for an additional distance $\frac{mg}{k}$ — thus displacing its center by an additional $\frac{mg}{2k}$. The center of mass of the top section is then at x-coordinate $\frac{l}{2}+\Delta x+\frac{mg}{2k}$ while that of the bottom section is

$$l+\Delta l+\frac{mg}{k}+\frac{l}{2}+\Delta x=\frac{3}{2}l+\frac{mg}{2k}+\frac{mg}{k}+\Delta x=\frac{3}{2}l+\frac{3mg}{2k}+\Delta x.$$

Remember that $\Delta l=\frac{mg}{2k}$ is the extension of the slinky under its own weight. The center of mass of the entire slinky of length $2l$ is then the average of the two coordinates as the two sections have equal masses.

$$x_{CM}=\frac{\frac{l}{2}+\Delta x+\frac{mg}{2k}+\frac{3}{2}l+\frac{3mg}{2k}+\Delta x}{2}=l+\Delta x+\frac{mg}{k}.$$

Furthermore, we know from scaling arguments that

$$x_{CM}=\frac{2l}{2}+4\Delta x=l+4\Delta x.$$

Equating these expressions,

$$\Delta x=\frac{mg}{3k}$$

$$\implies x_{CM}=\frac{l}{2}+\frac{mg}{3k}.$$

4.6.3 *Conservation of String*

Another important class of problems would be Atwood's machines. An Atwood's machine refers to a system of pulleys, strings and masses. For this chapter we will only consider the case where the pulleys and strings are massless, but this is not necessarily true in general.

Besides writing down the $F = ma$ equation[9] for each mass, we need to observe that the lengths of strings are "conserved." That is, the relative positions of the objects at the two ends of a string must obey a certain relationship as the string is inextensible. This observation will provide us with additional equations relating the accelerations of masses. Perhaps the next few examples will be somewhat enlightening.

Problem: Solve for the accelerations of m_1 and m_2, a_1 and a_2, and tensions in the set-up below.

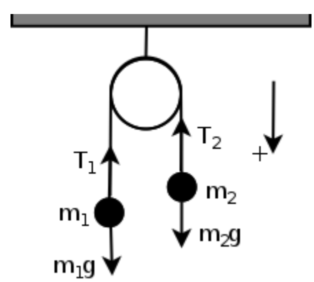


Figure 4.15: Atwood's machine 1

We first note that for the forces on the infinitesimal segment of string at the top of the pulley to be balanced (otherwise it will undergo infinite acceleration),

$$T_1 = T_2.$$

We shall just use T to denote tension thereafter. Actually, we can repeat this argument for every infinitesimal part of the massless string to conclude that the tension must be uniform throughout a continuous string segment so T is unambiguously the tension in the string connecting m_1 and m_2. Writing the $F = ma$ equations for both masses and defining downwards to be the

[9]We will not use vector notation as Atwood's machines are usually one-dimensional systems.

positive direction for all Atwood's machines in this chapter,

$$m_1 a_1 = m_1 g - T,$$
$$m_2 a_2 = m_2 g - T.$$

Lastly, applying the conservation of string, we must have:

$$a_1 = -a_2.$$

This is because if m_1 moves up a certain distance d, m_2 must also move down by d in order for the length of string to be "conserved". In solving the resultant set of simultaneous equations, it is usually expeditious to multiply the $F = ma$ equations by certain factors such that their sum becomes zero according to the conservation of string equation. Then, one can directly solve for the tension. For example, we can multiply the first $F = ma$ equation by m_2 and add it to the second $F = ma$ equation multiplied by m_1.

$$m_1 m_2 (a_1 + a_2) = 2 m_1 m_2 g - (m_1 + m_2) T.$$

Since $a_1 + a_2 = 0$,

$$T = \frac{2 m_1 m_2}{m_1 + m_2} g.$$

Subsequently, we substitute T back into the $F = ma$ equations to obtain

$$a_1 = \frac{m_1 - m_2}{m_1 + m_2} g,$$
$$a_2 = -a_1 = \frac{m_2 - m_1}{m_1 + m_2} g.$$

Problem: Solve for all tensions and accelerations in the two-layered system of three masses in Fig. 4.16.

We have already identified the relationship between tensions at different parts of the system. The tension in the string hanging over the bottom pulley is uniformly T due to the argument above. Then, the tension in the string hanging over the top pulley must be equal to $2T$ so that the massless bottom pulley-cum-bottom string system does not experience a net force.[10]

[10] The massless components can have non-zero accelerations but cannot experience non-zero net forces which will engender infinite acceleration. To further elaborate on the existence of an acceleration without a net force, Newton's second law for a massless particle yields $a = \frac{F}{m}$. If $F \to 0$ and $m \to 0$, a is generally indeterminate and depends on how F and m tend to zero. Therefore, a can generally undertake a non-zero value in spite of the lack of a net force.

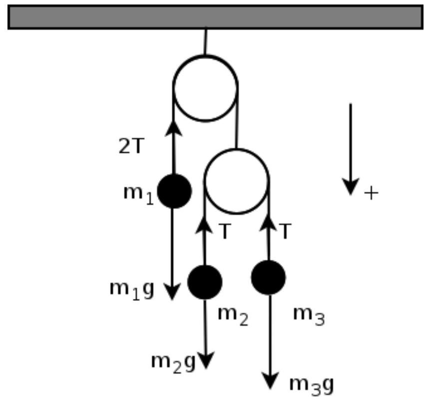


Figure 4.16: Atwood's machine 2

Therefore, we have learnt two conditions so far: tension is uniform throughout a massless, frictionless string while force balance must be attained at every massless pulley. Henceforth, we will directly indicate the tensions at different parts of the string, without explicit mention of these conditions, because the relationships are often trivial. However, if you are ever puzzled about why certain tensions are such and such, return to these two criteria and try to work them out yourself.

As always, we write our $F = ma$ equations for all three masses:

$$m_1 a_1 = m_1 g - 2T,$$

$$m_2 a_2 = m_2 g - T,$$

$$m_3 a_3 = m_3 g - T.$$

Lastly, the conservation of string equation is not that obvious in this case. It is in fact

$$a_1 = -\frac{a_2 + a_3}{2}.$$

This follows from the fact that the average vertical position of m_2 and m_3 moves the same distance as the bottom pulley which, in turn, moves the same distance as m_1. If this is still not obvious, we can split the motion of m_2 and m_3 into two components (Fig. 4.17). The first component would be due to the motion of the pulley they are clinging on to which accelerates at $-a_1$ due to the conservation of string. The second component can be seen as the acceleration a_f of the string connecting m_2 and m_3, around a pulley

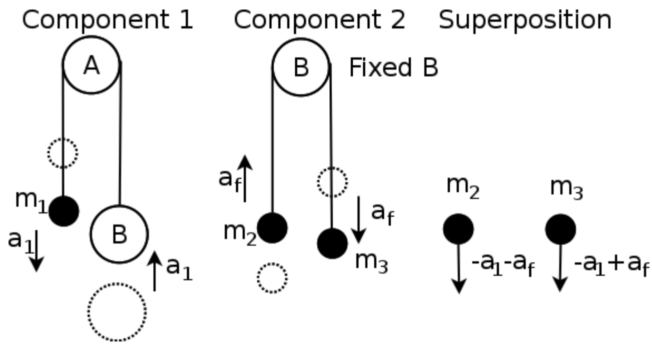


Figure 4.17: Different components of motion

that is at rest (as the motion of the pulley has been filtered out).

$$a_2 = -a_1 - a_f,$$
$$a_3 = -a_1 + a_f.$$

Thus,

$$a_1 = -\frac{a_2 + a_3}{2}.$$

Solving the four simultaneous equations (by multiplying the $F = ma$ equations by $2m_2m_3$, m_1m_3 and m_1m_2 respectively and adding them together), we obtain

$$T = \frac{4m_1m_2m_3}{4m_2m_3 + m_1(m_2 + m_3)}g,$$
$$a_1 = \frac{m_1(m_2 + m_3) - 4m_2m_3}{4m_2m_3 + m_1(m_2 + m_3)}g,$$
$$a_2 = \frac{4m_2m_3 + m_1(m_2 - 3m_3)}{4m_2m_3 + m_1(m_2 + m_3)}g,$$
$$a_3 = \frac{4m_2m_3 + m_1(m_3 - 3m_2)}{4m_2m_3 + m_1(m_2 + m_3)}g.$$

Problem: Let us now throw a movable pulley, which extends the breadth of a single layer, into the mix. Solve for all tensions and accelerations in the system depicted in Fig. 4.18.

Writing our $F = ma$ equations,

$$m_1a_1 = m_1g - T,$$
$$m_2a_2 = m_2g - 2T$$

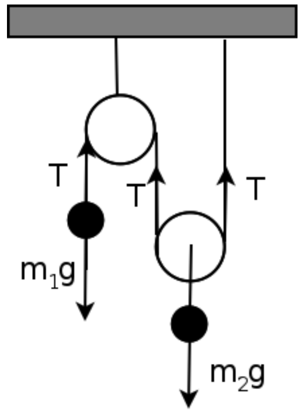


Figure 4.18: Atwood's machine 3

where the tension on m_2 is $2T$ for force balance at the right pulley. Observe that if the right pulley moves a distance of x downwards, m_1 must move $2x$ upwards in order for the length of the string to be conserved. Thus,

$$a_1 = -2a_2.$$

Solving (by multiplying the $F = ma$ equations by m_2 and $2m_1$ and adding),

$$T = \frac{3m_1m_2}{4m_1 + m_2}g,$$

$$a_1 = \frac{4m_1 - 2m_2}{4m_1 + m_2}g,$$

$$a_2 = \frac{m_2 - 2m_1}{4m_1 + m_2}g.$$

Weaving Constraints into Coordinates

Though the above derivations of the conservation of string equations provide vivid pictures of the physical situation, it is often easier to obtain the conservation of string equation by defining the coordinates of each mass while taking into account the length of each string. Then, the accelerations of the masses can be expressed in terms of fewer independent coordinates. In this process, it is convenient to assume that the lengths of all strings and the circumference of all pulleys are zero as they are inconsequential after subsequent differentiations.

Problem: Three identical masses m are connected in Fig. 4.19. Determine their accelerations.

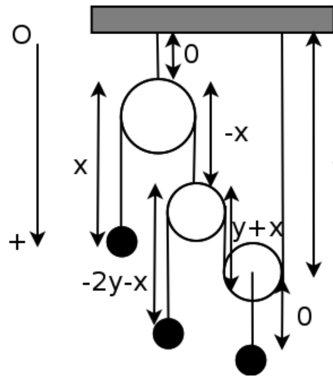


Figure 4.19: Atwood's machine 4

Number the masses in ascending order from left to right. Let the x-coordinate of the first mass be x (positive downwards). Then, the coordinates of the pulley connected to the string of the first mass is $0 - x = -x$ as the length of the string connecting them must be zero. Let the x-coordinate of the third mass, and thus the pulley it is connected to, be y. Then, the segment between the second and third pulleys is of length $y - (-x) = y + x$. Lastly, the length of the segment between the top of the second pulley and the second mass is then $0 - y - (y + x) = -2y - x$ as the length of the string wrapping over the second and third pulleys is zero. The corresponding coordinates of the masses are then x, $-x + (-2y - x) = -2x - 2y$ and y. Hence,

$$a_1 = \ddot{x},$$
$$a_2 = -2\ddot{x} - 2\ddot{y},$$
$$a_3 = \ddot{y}.$$

Evidently, the conservation of string equation is

$$2a_1 + a_2 + 2a_3 = 0.$$

Let the tension on the second mass be T. Then the tensions on the first and third masses are both $2T$. $F = ma$ yields

$$ma_1 = mg - 2T,$$
$$ma_2 = mg - T,$$
$$ma_3 = mg - 2T.$$

Solving,

$$T = \frac{5}{9}mg,$$
$$a_1 = -\frac{1}{9}g,$$
$$a_2 = \frac{4}{9}g,$$
$$a_3 = -\frac{1}{9}g.$$

4.6.4 *Remaining on an Inclined Plane*

In certain problems, objects are constrained to remain on a surface such as an inclined plane. Then, the acceleration of the object relative to the surface must satisfy a certain relationship. Consider the following problem.

Problem: A block of mass m lies on a frictionless plane of mass M and angle of inclination θ. If you maintain the horizontal acceleration of the plane at A, determine the acceleration of the block and the normal force on the mass due to the plane.

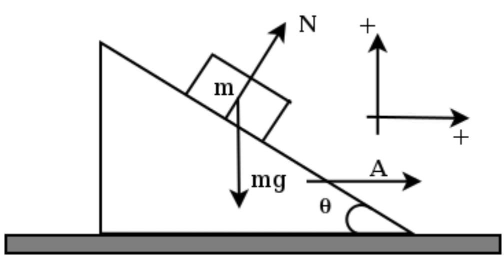


Figure 4.20: Block on accelerating inclined plane

As usual, writing our $F = ma$ equations in the horizontal and vertical directions,

$$N \sin\theta = ma_x,$$
$$N\cos\theta - mg = ma_y.$$

We have three variables (N, a_x and a_y) and two equations — hence we still require one more. This comes from the astute observation that the block must remain in contact with the plane. Under this constraint,

$$\frac{a_y}{a_x - A} = -\tan\theta.$$

This constraint simply means that the magnitude of the relative acceleration between the mass and plane in the vertical direction must be $\tan\theta$ times that in the horizontal direction, for the block to remain on the plane. The negative sign stems from the fact that if the block moves upwards relative to the plane, it must move leftwards relative to the plane. Solving this set of equations yields

$$\begin{aligned} a_x &= g\sin\theta\cos\theta + A\sin^2\theta, \\ a_y &= A\sin\theta\cos\theta - g\sin^2\theta, \\ N &= mg\cos\theta + mA\sin\theta. \end{aligned}$$

Another approach to solving this system would be to define two independent coordinates. Let the top-left tip of the plane be at coordinates $(x, 0)$. Then, define s to be the distance of the block from the top-left tip of the plane. The coordinates of the block are then $(x + s\cos\theta, -s\sin\theta)$. Hence,

$$\begin{aligned} a_x &= \ddot{x} + \ddot{s}\cos\theta = A + \ddot{s}\cos\theta, \\ a_y &= -\ddot{s}\sin\theta, \\ N\sin\theta &= m(A + \ddot{s}\cos\theta), \\ N\cos\theta - mg &= -m\ddot{s}\sin\theta. \end{aligned}$$

Solving this new set of equations will yield the same result as above, because the condition for the block to remain on the plane has been subtly included in the definition of the coordinates. To show this,

$$\frac{a_y}{a_x - A} = -\frac{\ddot{s}\sin\theta}{\ddot{s}\cos\theta + \ddot{x} - A} = -\tan\theta.$$

4.6.5 *Polar Coordinates*

Finally, we shall practise solving some systems in polar coordinates. An important constraint in polar coordinates is that of circular motion. It was derived in Chapter 3 that for an object to remain in a circle of constant radius

r, it must experience an instantaneous centripetal acceleration, directed radially inwards, that is equal to

$$a_r = -\frac{mv^2}{r},$$

where v is its instantaneous tangential velocity. Therefore, there must be a net centripetal force on the object which obeys

$$\sum \boldsymbol{F}_r = -\frac{mv^2}{r}\hat{\boldsymbol{r}} = -mr\dot{\theta}^2\hat{\boldsymbol{r}},$$

by Eq. (4.10) since $\ddot{r} = 0$.

Problem: Consider the following conical pendulum of length l which undergoes uniform circular motion at a constant vertical height with an angular velocity ω. Determine the range of ω for which the pendulum is able to maintain such a motion at $\theta > 0$. What if ω is smaller than the lower bound of this range?

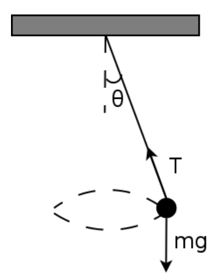


Figure 4.21: Conical pendulum

For the pendulum to stay at the same vertical height, the net force in the vertical direction must be zero.

$$T\cos\theta = mg.$$

Furthermore, the pendulum undergoes uniform circular motion with a radius of rotation $l\sin\theta$. Therefore, the radial component of tension must provide the required centripetal force.

$$T\sin\theta = ml\sin\theta\omega^2.$$

As we are considering positions at which $\theta > 0$, we can simply cancel the $\sin\theta$ terms.

$$T = ml\omega^2$$
$$mg = ml\omega^2\cos\theta$$
$$\cos\theta = \frac{g}{l\omega^2}.$$

As $|\cos\theta| \leq 1$,

$$\omega \geq \sqrt{\frac{l}{g}}$$

for the required motion to be possible. When $\omega < \sqrt{\frac{l}{g}}$, the math breaks down when we cancel the $\sin\theta$'s as $\theta = 0$ in this case (the pendulum cannot sustain circular motion). Thus, θ as a function of ω is in fact

$$\theta = \begin{cases} \cos^{-1}\frac{g}{l\omega^2} & \text{for } \omega \geq \sqrt{\frac{l}{g}} \\ 0 \quad \text{for } \omega < \sqrt{\frac{l}{g}}. \end{cases}$$

Let us now consider a more general application of $F = ma$ in polar coordinates.

Problem: Two masses are connected via a string through a hole on a horizontal table. Mass m lies on the plane and is currently moving with a radial velocity v_r, with the positive direction taken to be radially outwards, and angular velocity ω at a radius r away from the hole. Determine the instantaneous acceleration of the mass M below the table and the instantaneous angular acceleration of mass m.

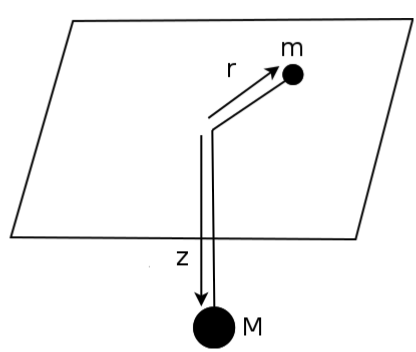


Figure 4.22: Two masses

Let the tension in the string be T. Applying Newton's laws to mass m in polar coordinates in the plane of the table,

$$-T = m\ddot{r} - mr\dot{\theta}^2,$$
$$0 = mr\ddot{\theta} + 2m\dot{r}\dot{\theta}.$$

At this instant, $\dot{r} = v_r$ and $\dot{\theta} = \omega$.

$$-T = m\ddot{r} - mr\omega^2,$$
$$0 = mr\ddot{\theta} + 2mv_r\omega.$$

For mass M,

$$Mg - T = M\ddot{z}.$$

Furthermore, the conservation of string requires $\ddot{r} = -\ddot{z}$. Solving,

$$\ddot{z} = \frac{Mg - mr\omega^2}{m + M},$$
$$\ddot{\theta} = -\frac{2v_r\omega}{r}.$$

Note that all of these quantities are instantaneous. If we observe the second result, we can see that if the top mass moves radially outwards, its angular velocity will decrease at the next instance and vice-versa.[11]

4.6.6 *Rigid Body Constraint*

Another common constraint is the rigid body criterion where particles have to maintain fixed relative distances with respect to one another. A usual set-up consists of discrete particles connected by massless, rigid rods (if the rigid body is a continuous body, it is better described by other methods). As a rigid rod often only experiences forces at its two ends, the tension due to the rod can only be parallel to itself in order for the forces and torques on it to be balanced (the lines of the equal and opposite forces, which nullify each other, can only coincide and thus lie along the rod — otherwise, taking moments about one end would result in a non-zero net torque). This observation, coupled with the kinematics of rigid body motion discussed in Chapter 3, is crucial in solving such problems.

Problem: Three masses m_1, m_2 and m_3 are connected via massless, rigid rods of lengths l_1, l_2 and l_3 to a common massless center O. The connection

[11] This is expected as the angular momentum of this system about the hole must be conserved.

at O is fixed such that adjacent rods subtend constant angles $\theta_1 > 0$, $\theta_2 > 0$ and $\theta_3 > 0$. The set-up lies on a frictionless, horizontal table and the masses are initially given velocities v_1, v_2 and v_3, perpendicular to their respective rods and in the clockwise direction. Determine the instantaneous acceleration of each mass.

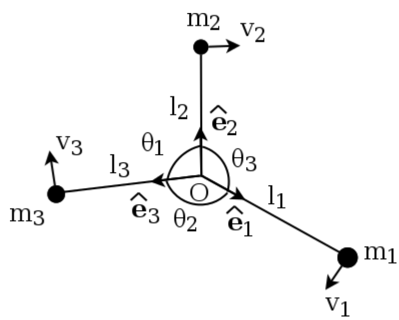


Figure 4.23: Three masses and rigid rods

Firstly, we shall show that the velocity of O is zero. In this problem, we define unit vectors $\hat{\boldsymbol{e}}_1$, $\hat{\boldsymbol{e}}_2$ and $\hat{\boldsymbol{e}}_3$ which are parallel to the respective rods, emanating from the center O (which also functions as the origin). Let the velocity of the center be $\boldsymbol{v}_O$. The rigid body constraint requires that there be no relative velocity between two ends of a rod, along the rod. Expressing this condition vectorially for all three rods,

$$(\boldsymbol{v}_O - \boldsymbol{v}_i) \cdot \hat{\boldsymbol{e}}_i = 0,$$

for all $1 \leq i \leq 3$, where $\boldsymbol{v}_i$ is the velocity of the ith mass. Since $\boldsymbol{v}_i \cdot \hat{\boldsymbol{e}}_i = 0$,

$$\boldsymbol{v}_O \cdot \hat{\boldsymbol{e}}_i = 0$$

for all i. Since the $\hat{\boldsymbol{e}}_i$'s are in different directions,[12] the only possible way for the above to be satisfied is for $\boldsymbol{v}_O = \boldsymbol{0}$. Moving on, let the tensions exerted on the massless center O by the rods be $\boldsymbol{T}_1$, $\boldsymbol{T}_2$ and $\boldsymbol{T}_3$ respectively. The

[12]The more rigorous explanation is that any pair of $\hat{\boldsymbol{e}}_i$'s would function as a set of basis vectors for two-dimensional space. Since any vector in the plane of the table can be expressed as a linear combination of the basis vectors, a vector whose dot products with both basis vectors are zero can only be the null vector.

forces on O must be balanced — lest it experiences infinite acceleration.

$$\boldsymbol{T}_1 + \boldsymbol{T}_2 + \boldsymbol{T}_3 = \boldsymbol{0}.$$

That is, the three tension vectors form a force triangle depicted below. The relative angles between these tensions are given by the fact that they are directed along the respective rods.

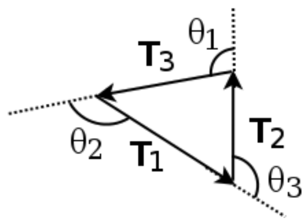


Figure 4.24: Force triangle

By the sine rule,

$$\frac{T_1}{\sin\theta_1} = \frac{T_2}{\sin\theta_2} = \frac{T_3}{\sin\theta_3} = T,$$

for some new variable T. Furthermore, since $\boldsymbol{T}_i = T\sin\theta_i \boldsymbol{e}_i$,

$$\sum_{i=1}^{3} T\sin\theta_i \boldsymbol{e}_i = \boldsymbol{0}$$

$$\implies \sum_{i=1}^{3} \sin\theta_i \boldsymbol{e}_i = \boldsymbol{0}.$$

Now, denote the acceleration of the center O as $\boldsymbol{a}_O$ (remember that even though there is no net force, the massless connection can have an acceleration). Based on the rigid body constraint, the relative acceleration between two ends of a rod, along the rod, must reflect the centripetal acceleration associated with the relative tangential velocity (in this case, it is simply the $\boldsymbol{v}_i$'s since the center O is stationary). Since the acceleration of m_i is $\boldsymbol{a}_i = -\frac{T_i}{m_i}\hat{\boldsymbol{e}}_i = -\frac{T\sin\theta_i}{m_i}\hat{\boldsymbol{e}}_i$, the above condition requires

$$\frac{T\sin\theta_i}{m_i} + \boldsymbol{a}_O \cdot \hat{\boldsymbol{e}}_i = \frac{v_i^2}{l_i}$$

for all $1 \leq i \leq 3$. Multiplying the ith equation by $\sin\theta_i$ and summing everything,

$$\sum_{i=1}^{3} \frac{T\sin^2\theta_i}{m_i} = \sum_{i=1}^{3} \frac{v_i^2 \sin\theta_i}{l_i},$$

as $\boldsymbol{a}_O \cdot \left(\sum_{i=1}^{3} \sin\theta_i \hat{\boldsymbol{e}}_i\right) = \boldsymbol{a}_O \cdot \boldsymbol{0} = 0$.

$$T = \frac{\sum_{i=1}^{3} \frac{v_i^2 \sin\theta_i}{l_i}}{\sum_{i=1}^{3} \frac{\sin^2\theta_i}{m_i}}.$$

Finally, the acceleration of the ith mass is simply

$$\boldsymbol{a}_i = -\frac{T\sin\theta_i}{m_i}\hat{\boldsymbol{e}}_i = -\frac{\sum_{i=1}^{3} \frac{v_i^2 \sin\theta_i}{l_i}}{\sum_{i=1}^{3} \frac{\sin^2\theta_i}{m_i}} \cdot \frac{\sin\theta_i}{m_i}\hat{\boldsymbol{e}}_i.$$

4.7 Systems with Variable Amounts of Moving Mass

In mechanics, a system is not defined as a physical region demarcated by a boundary. This incorrect notion implies that if certain particles enter or exit a certain region, we include and exclude them in our system respectively. Instead, a system is defined as a predetermined set of particles which are tracked by us thereafter. Therefore, it is deceptive to write the following equation for a rigid-body system that is purely translating, in hopes that the $\frac{dm}{dt}$ term represents a physical increase or decrease in mass inside the rigid-body system.

$$\sum \boldsymbol{F} = \frac{d\boldsymbol{P}}{dt} = \frac{d(M\boldsymbol{v}_{CM})}{dt} = \frac{dM}{dt}\boldsymbol{v}_{CM} + M\boldsymbol{a}_{CM}.$$

The above equation implies that at one time, we are considering a certain system of particles and at another time, we are considering another completely different system! Instead, we should define an all-encompassing system with different masses moving at different velocities. Then we can find the total momentum of this all-encompassing system as a function of time $\boldsymbol{p}(t)$[13] and take its time derivative to be related to the net external force on this system. Consider the following examples.

Problem: An empty box of sand is initially moving at a horizontal velocity v on a frictionless ground. You proceed to add more initially-stationary sand into the box at a rate of σ (mass added per unit time). What is the horizontal force required to keep the box traveling at a constant velocity?

We consider the system of the cart (with the sand inside) and all of the sand that may or may not have been added to the cart. After time t, the

[13] We shall use the smaller-case 'p' as we will only be considering a single system.

total horizontal momentum of this system would be

$$p = (m_0 + \sigma t)v,$$

where m_0 is the initial mass of the cart. Notice that the net external force on this combined system in the horizontal direction is simply that which you exert on the cart. Hence,

$$F = \frac{dp}{dt} = \sigma v.$$

Incidentally, if we applied the equation

$$F = ma + \frac{dm}{dt}v,$$

we would obtain the same result as $a = 0$ and $\frac{dm}{dt} = \sigma$. However, the physical meaning of the above equation is wrong — as shall be illustrated by the following example with a subtle difference.

Problem: An empty box of sand is initially moving at velocity v. You hold sand in your hand of total initial mass M and move it at velocity u in the same direction as the box. You then release sand (which still has velocity u immediately after escaping from your hand) into the box at a rate of σ. What is the force required to keep the box traveling at a constant velocity?

Applying $F = ma + \frac{dm}{dt}v$ would give the same result σv which is evidently incorrect. Instead, we should consider the cart and the sand to be a whole system. After time t, the masses of the cart and remaining sand in your hand are $m_0 + \sigma t$ and $M - \sigma t$ respectively. Hence, the total horizontal momentum of the combined system as a function of time is

$$p = (m_0 + \sigma t)v + (M - \sigma t)u.$$

Since the net external force in the horizontal direction on this combined system is that by you on the cart,

$$F = \frac{dp}{dt} = \sigma(v - u).$$

Lastly, let us consider an extreme example where the $\frac{dm}{dt}$ term results in an utterly unreasonable answer.

Problem: A box of sand initially moving at a velocity v leaks sand at a rate of σ (mass lost per unit time). There is no internal friction between the sand and the cart such that the ejected sand still travels at velocity v. What is the force required to keep the box traveling at a constant velocity?

Applying $F = ma + \frac{dm}{dt}v$ would yield $-\sigma v$ which implies that an opposing force needs to be applied to the cart to prevent it from accelerating! If we instead consider the box and sand as a combined system, we will observe that the total horizontal momentum of this system does not vary with time (assuming that no other horizontal forces act on the ejected sand) and hence conclude that no force is in fact required!

Another perspective to the above problems involves splitting the process into two parts. First, the sand is added or lost, which may or may not result in a change of the cart's velocity. Next, the force required to keep it moving at a constant velocity is applied. For the first problem, the first event results in a decrease in the cart's velocity by the conservation of momentum and thus requires a force to keep the cart moving at a constant velocity during the second event. For the second problem, the first event results in a smaller decrease in the cart's velocity as the sand was already moving and thus requires a smaller force during the second event. Lastly, the first event in the third problem does not result in a change in the cart's velocity and thus a force on the cart is unnecessary. This approach — which entails the consideration of the effects of adding an infinitesimal amount of mass and applying a force over an infinitesimal duration of time — will be introduced in Chapter 6, where the impulse-momentum theorem will also be discussed.

It is worthy to note that the latter method has a salient advantage over that presented in this chapter as it describes the situation more precisely. By introducing an all-encompassing system, we lose the specificity of our analysis. For example, in the second problem, once some sand has been dropped, the motion of the cart should be independent of the motion of the sand still remaining in our hand (we could accelerate it, for example). In the third problem, whether the ejected sand experiences any force or not should have no effect on the motion of the cart. Ultimately, interactions should only be local and this principle of locality is a sacrosanct pillar of physics — for which casting a wider net only blurs. However, we can conversely say that since interactions should only be local, we can tweak the parts of a system that should have no direct influence on a certain component of interest in any manner and the motion of that component should still be the same as before — justifying our assumptions in the previous problems.

For now, let us practise deriving p and finding $\frac{dp}{dt}$ for the following systems.

Problem: A uniform, straightened chain with linear mass density λ and length l hangs vertically at rest with one tip on the surface of a weighing scale, and is then released. Assuming that the parts of the chain that hit the scale immediately come to rest, find the reading on the weighing scale as a function of the vertical distance that the top end of the chain has fallen, x. The bend at the surface of the scale is small.[14]

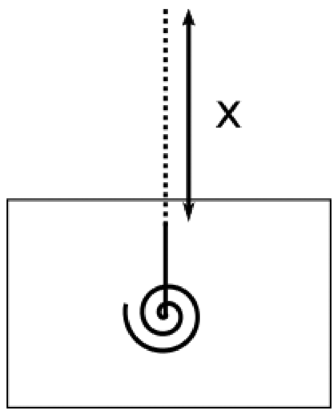


Figure 4.25: Falling chain

Define the positive vertical direction to be downwards. The first important observation to make is that the tension in the chain must be zero as the tensions on the massless bend are along the vertical and horizontal directions and thus cannot balance each other. The part of the chain that has yet to collide with the scale is hence free falling and its velocity can be obtained from the kinematics equation below.

$$v^2 = 2gx$$
$$v = \sqrt{2gx}.$$

The length of the chain that is still moving is $(l - x)$. The total momentum of the entire chain, including the stationary parts that are resting on the scale, is

$$p = \lambda(l - x)\dot{x}.$$

Consequently,

$$\frac{dp}{dt} = -\lambda\dot{x}^2 + \lambda(l - x)\ddot{x}.$$

[14] Small, in the sense that the length of the bend is much smaller than the separation between infinitesimal masses in the rope, such that the bend is essentially massless.

Substituting $\dot{x} = v = \sqrt{2gx}$ and $\ddot{x} = g$,

$$\frac{dp}{dt} = -3\lambda gx + \lambda gl.$$

The net external force on the chain is due to its weight and the normal force due to the scale.

$$\lambda gl - N = -3\lambda gx + \lambda gl$$
$$N = 3\lambda gx.$$

Note that at $x = l$, the normal force abruptly decreases from $3\lambda gl$ to λgl (the total weight of the chain). This decrease originates from the depletion of parts of the chain that crash into the scale (the exact reduction of $2\lambda gl$ due to this reason is best examined via the method in Chapter 6).

Problem: A long, uniform chain of linear mass density λ is twined into an initially-stationary heap with an infinitesimal segment of one end hanging from a hole on a horizontal, frictionless table. Assuming that there is no internal friction between segments of the chain and that the only moving part of the chain is that below the table, determine the velocity of the moving part of the chain, $\dot{x}$, as a function of the length x that has fallen below the table and thus, $x(t)$.

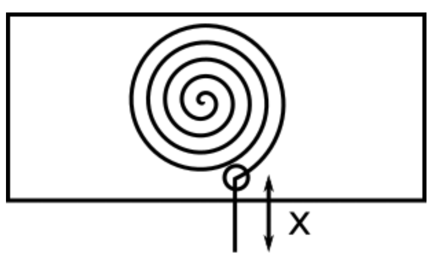


Figure 4.26: Falling chain from heap

The momentum of the entire chain, which hinges on the moving part, is $p = \lambda x\dot{x}$. The net external force on the entire chain in the vertical direction is λgx (weight of the hanging segment) as the normal force on the segment of the chain above the table and its weight should exactly cancel for its vertical momentum to remain zero. Furthermore, there must be no tension in the chain as the chain on the frictionless table cannot experience a net horizontal tension force due to the chain segment in the hole.

$$\lambda gx = \frac{dp}{dt} = \lambda\dot{x}^2 + \lambda x\ddot{x}.$$

Using the trick $\ddot{x} = \frac{1}{2}\frac{d\dot{x}^2}{dx}$ and simplifying,

$$\frac{d\dot{x}^2}{dx} + \frac{2}{x}\dot{x}^2 = 2g.$$

Multiplying the above by the appropriate integrating factor x^2,

$$\begin{aligned}
x^2\frac{d\dot{x}^2}{dx} + 2x\dot{x}^2 &= 2gx^2 \\
\frac{d(x^2\dot{x}^2)}{dx} &= 2gx^2 \\
\int_0^{x^2\dot{x}^2} d(x^2\dot{x}^2) &= \int_0^x 2gx^2 dx \\
x^2\dot{x}^2 &= \frac{2}{3}gx^3 \\
\dot{x} &= \sqrt{\frac{2}{3}gx},
\end{aligned}$$

where the negative solution has been rejected because the velocity of the falling part can only be downwards (positive). This follows from the fact that the net force on the system is downwards. Then,

$$\begin{aligned}
\int_0^x \frac{1}{\sqrt{x}}dx &= \int_0^t \sqrt{\frac{2}{3}}gdt \\
2\sqrt{x} &= \sqrt{\frac{2}{3}}gt \\
x &= \frac{1}{6}gt^2.
\end{aligned}$$

Problem: A uniform, straightened chain of total mass m and length l is initially stationary on a horizontal, frictionless table. Initially, a segment of infinitesimal length, at one end of the chain, passes through a hole on the table and hangs vertically. Assuming that segments of the chain, at no point in time, “overshoot” the hole, determine the velocity of the chain, $\dot{x}$, as a function of the length x that has fallen below the table and thus, $x(t)$. You will notice that there is an error in $x(t)$. What is the reason behind this error and why did it not occur in the previous problem?

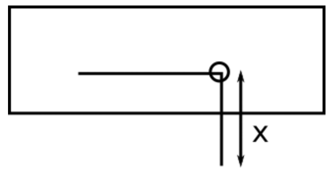


Figure 4.27: Falling linear chain

The tension T located at the chain segment within the hole changes the horizontal momentum of the chain segment on the table, $\frac{m}{l}(l-x)\dot{x}$.

$$T = \frac{d\left(\frac{m}{l}(l-x)\dot{x}\right)}{dt}.$$

On the other hand, the weight of the hanging section $\frac{mg}{l}x$, minus this tension, is responsible for changing the vertical momentum of the hanging part of the chain, $\frac{m}{l}x\dot{x}$.

$$\frac{mg}{l}x - T = \frac{d\left(\frac{m}{l}x\dot{x}\right)}{dt}.$$

Adding the two equations above,

$$\frac{mg}{l}x = \frac{d(m\dot{x})}{dt} = m\ddot{x}.$$

Using $\ddot{x} = \dot{x}\frac{d\dot{x}}{dx}$,

$$\int_0^{\dot{x}} \dot{x}d\dot{x} = \int_0^x \frac{g}{l}xdx$$
$$\frac{\dot{x}^2}{2} = \frac{gx^2}{2l}$$
$$\dot{x} = \sqrt{\frac{g}{l}}x.$$

The positive solution for $\dot{x}$ is chosen as x is initially slightly positive and $\ddot{x}$ is proportional to x throughout the motion. Now, if we attempt to separate variables and integrate,

$$\int_{0^+}^x \frac{1}{x}dx = \int_0^t \sqrt{\frac{g}{l}}dt$$
$$\ln x = \sqrt{\frac{g}{l}}t + \ln 0^+.$$

A pesky $\ln 0^+$ term appears out of nowhere! The reason behind this term is the infinite time required for the chain to fall through the hole from an initially negligible length to an appreciable length (hence causing $\ln x$ to tend to negative infinity). This defect is present in this problem as the chain is not slack such that the tension in the chain is equal to the weight of the infinitesimal hanging segment at the start to prevent infinite acceleration. This small tension pulls the portion of the chain on the table into the hole (giving it momentum in the process). Since the latter portion is massive, it naturally takes an infinite amount of time to pull it into the hole by a significant amount (and to impart it with non-negligible momentum). In the previous problem, the entire chain was slack such that no force is required to drag the heap lying on the table into the hole (the heap simply readjusts itself while maintaining the position of its center of mass). Furthermore, initially, the weight of the infinitesimal segment hanging below the hole directly contributes to its rate of change of momentum — without any impediment from tension. Once it weathers the initial storm and attains a non-negligible hanging length, the hanging portion is able to perpetuate a self-sustaining cycle where a larger x leads to a larger $\dot{x}$ (this also occurs for the straight chain but it fails at the first hurdle). From the perspective of momentum, the weight of the (initially negligible) hanging portion only has to increase its own momentum and not that of the entire chain and hence can attain a substantial velocity.

In spite of all of these remarks, the expression for $\dot{x}(x)$ in this problem is still valid — it just takes an infinite amount of time to reach a non-negligible x.

Problems

Center of Mass

1. *Regular Pentagon**

Determine the center of mass of a uniform, regular pentagon with edge length l.

2. *A Strange Rod**

Find the center of mass of a one-dimensional rod of length l, whose linear mass density at a point is $\lambda = \frac{k}{x+l}$ where k is a constant and x is the distance from the left end of the rod to that point.

3. *Cone***

Determine the vertical height below the vertex of a uniform cone with base area A and height h, at which the center of mass lies.

4. *Spherical Cap***

Show that the center of mass of a spherical cap of height h, cut from a uniform sphere of radius R, is

$$\frac{3(2R-h)^2}{4(3R-h)}$$

above the center of the original sphere. Note that a spherical cap is obtained from slicing a sphere with a plane.

5. *Triangle***

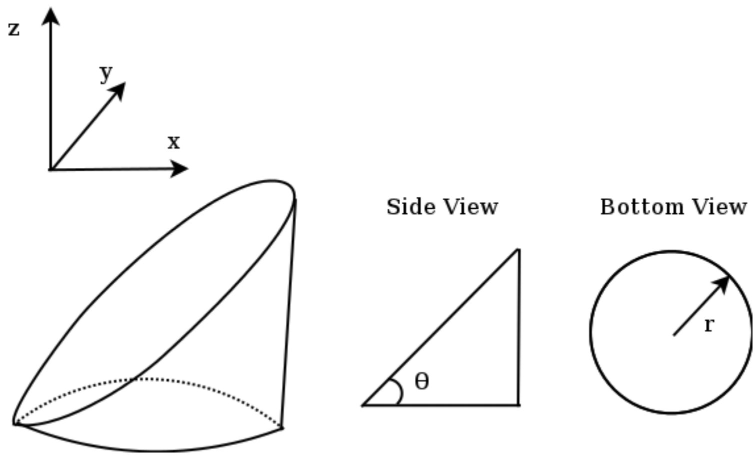
Express the coordinates of the center of mass of a general uniform triangle in terms of the coordinates of its three vertices. Hint: try to apply the result for a right-angled triangle.

A plated triangle is formed as follows. Take a uniform triangle of surface mass density σ — whose vertices have coordinates (x_1, y_1), (x_2, y_2) and (x_3, y_3) — and connect the midpoints of the edges. Fill the new triangle formed with a plating with surface mass density σ and repeat this process with the new triangle and all subsequent triangles. Determine the center of mass of this plated triangle.

As an unrelated question, prove geometrically that the center of mass of a uniform triangle should be the point of intersection of the three medians (you do not need to prove that the medians are concurrent).

6. *Cylindrical Segment*******

Find the position of the center of mass of the following cylindrical segment, whose base is a circle of radius r, that has a uniform mass density. It is obtained from slicing a cylinder with a plane.

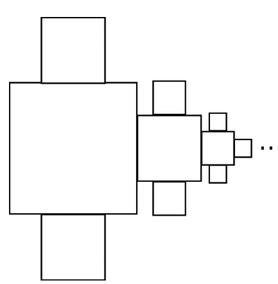


7. *Constant Ratio*******

Let a rod of length l and mass M be located along the x-axis with ends at $x = 0$ and $x = l$. Now, this rod has a special property such that if we make a cut at any arbitrary $x = y$ coordinate and consider the remaining rod between $x = 0$ and $x = y$, the center of mass of this remaining rod will be located at ky, where k is a constant. Determine $\lambda(x)$, the linear mass density of the rod as a function of the x-coordinate x.

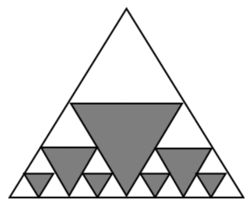
8. *Square Fractal*******

Determine the center of mass of the square fractal below if its surface mass density is uniform. The largest square has length l, after which each successive square has half the length of its predecessor. Note that the fractal is only "propagating" in one direction.



9. *Triangle Fractal****

Consider a triangle fractal that is obtained by filling up an equilateral triangle in a equilateral triangle of length l with mass (which occupies a quarter of the area of the original triangle) to produce three empty equilateral triangles of length $\frac{l}{2}$. The bottom two triangles then undergo the same procedure again. This process is repeated for all subsequent triangles indefinitely. Determine the vertical distance between the center of mass of this fractal and the bottom edge of the original triangle in the figure below.



Systems with No Constraints

10. *Walking on a Plank**

Suppose that a wooden plank of length l and mass M lies on a frictionless horizontal ground. Consider the one-dimensional problem where a person of mass m starts from one end of the plank and moves to the other end of the plank. Determine the horizontal displacement of the plank if both the plank and person were initially stationary.

Now, suppose that a massless ant initially rests at one end of the plank. The ant has a stock of snowballs of total mass m. If the ant begins to throw snowballs at velocity v and they stick to a massless wall at the other end of the plank, determine the horizontal displacement of the plank after all snowballs have hit the wall. The ant and plank were both stationary initially.

11. *Pulling a Block**

You and a block of metal are initially motionless on a frictionless horizontal ground, separated by a distance l. Your mass is m while that of the block is M. You begin to pull the block at a constant force F via a massless string. Determine the time t at which you collide with the block.

12. *Falling Slinky**

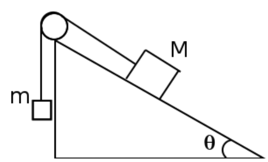
A slinky with negligible relaxed length, mass m and spring constant k is held vertically by its top in mid-air. The top of the slinky is released such that it begins to fall down and collide with other segments of the slinky. If the segments that have yet to collide with the top of the slinky are observed to remain still, determine the time from the release of the slinky that the bottom of the slinky begins moving. Repeat your calculations if an additional mass m is hung to the bottom of the slinky at the start.

Systems with Constraints

For the following problems, assume that there is no friction between all surfaces and that strings and pulleys have negligible masses, unless otherwise stated.

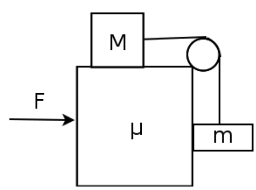
13. *Atwood's Machine 1**

Find the accelerations of masses m and M if the ramp does not move. Friction exists between the ramp and M, with a static coefficient μ_s and kinetic coefficient $\mu_k < \mu_s$.



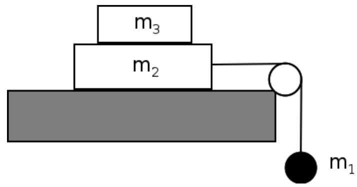
14. *Atwood's Machine 2**

Find the force F required to prevent any relative motion of m, M and μ.



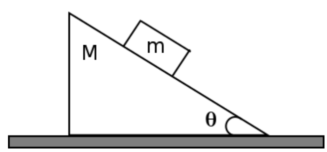
15. *Traveling Together***

Determine the maximum value of m_1 in the set-up on the next page if masses m_2 and m_3 remain stationary with respect to each other. The coefficient of kinetic friction between m_2 and the table is μ_k while the coefficient of static friction between m_2 and m_3 is μ_s. Assume that m_2 moves relative to the table.



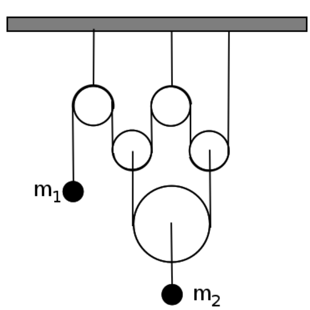
16. ***Sliding down a Plane*****

A block of mass m is held motionless on a frictionless plane of mass M and angle of inclination θ. There is no friction between the inclined plane and the ground. The block is released. What is the horizontal acceleration of the plane?



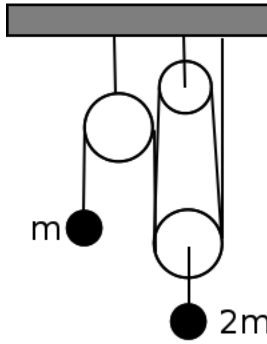
17. ***Atwood's Machine 3*****

Find all tensions and the accelerations of the masses in the set-up below.



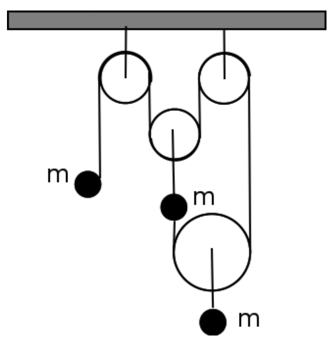
18. ***Atwood's Machine 4*****

Determine the accelerations of the masses m and $2m$. For the two pulleys on the next page, a continuous string wraps once around the bottom pulley, once around the top pulley and one last time around the bottom pulley before it is connected to the ceiling.



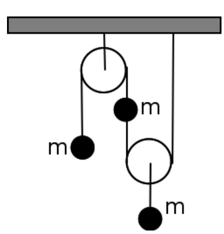
19. ***Atwood's Machine 5*****

Determine the accelerations of all masses in the set-up below. Try to obtain the conservation of string equation by visualizing the physical movement of strings.



20. ***Atwood's Machine 6*****

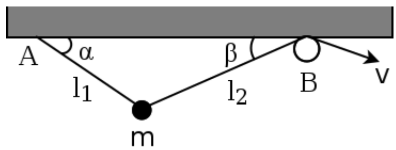
Determine the accelerations of all masses in the set-up below.



21. ***Pulling a Mass*****

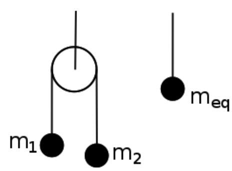
A ball of mass m is attached to two strings — one is of length l_1 and is attached to a fixed pivot A, while the other is wrapped around a fixed pulley

B and dragged along at a constant velocity v. The instantaneous length of the segment between the mass and B is l_2. As a function of angles α and β labelled in the diagram below, determine the tension on the ball due to the second string.



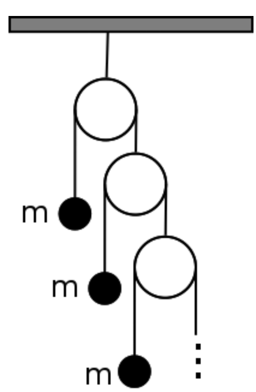
22. *Equivalent Mass**

Consider two masses m_1 and m_2 connected by a pulley. Suppose that we are able to replace this set-up with an equivalent one that comprises a single mass m_{eq} connected to a string. Determine m_{eq}. Hint: if the set-ups are equivalent, the behavior of any system connected to them should be the same.



23. *Infinite Atwood Machine***

An infinite number of identical masses are arranged as shown below. Determine the acceleration of all masses. Assume that the set-up comprises N masses with the Nth mass replacing what would have been the Nth pulley. Then, take $N \to \infty$. What would happen if the Nth mass were a massless pulley instead? The result of the above question may be useful. (Adapted from "Introduction to Classical Mechanics.")

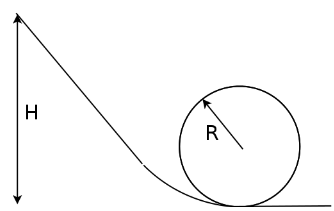


Polar Coordinates/Circular Motion

Note that though the conservation of energy has not been formally introduced yet, it is extremely useful for many of the problems below and will be adopted in their solutions.

24. *Roller Coaster**

A roller coaster slides down a slope. Find the minimum height, H, necessary for the roller coaster to undergo circular motion of radius R at the bottom of the slope without losing contact with the surface.

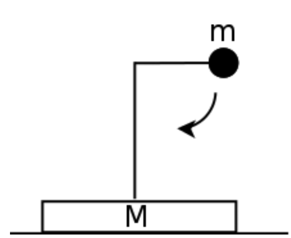


25. *Falling off a Circle***

A particle, of mass m, initially lies motionless on the top of a circle of radius R. It is then given a slight push and proceeds to move along the surface of the circle. At what angle, θ_c, measured with respect to the vertical axis, will the particle lose contact with the circle?

26. *Stationary Stand***

A stand is comprised of a base of mass M connected to a massless pole. A mass m is attached to the tip of the pole via a massless string of length l and is initially held motionless at the same horizontal level as the tip of the pole. The mass m is then released. Determine the minimum coefficient of static friction μ between the stand and the table if the stand does not translate while mass m undergoes circular motion about the tip of the pole thereafter.



27. *Rotating Rod***

Consider a uniform rod of linear mass density λ and length l on a horizontal table in polar coordinates, with the origin defined at one of its ends. Supposing that the rod rotates about the origin at a constant angular velocity ω, determine the tension as a function of radial coordinate $T(r)$ if the end at $r = 0$ is free (the end at $r = l$ could be connected to a rotating cylinder, for example) and if the end at $r = l$ is free (the end at $r = 0$ could be skewered and rotated, for instance).

28. *Rotating Chain***

A chain of uniform linear mass density λ takes the form of a circle of radius R and is wrapped around a frictionless cone with half angle α. If the chain is able to rotate about the symmetrical axis of the cone at constant angular velocity ω while maintaining its shape of a circle of radius R, determine the tension in the chain T.

Systems with Varying Amounts of Moving Mass

29. *Sweeping Pan***

Consider an initially stationary sweeping pan of width l and initial mass M on a horizontal, frictionless table. There is dust uniformly distributed over the entire table with surface mass density σ. Determine $F(t)$, the force required to push the pan such that it accelerates with a constant acceleration a.

30. *Holding a Rope***

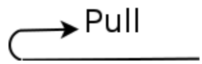
Take a uniform rope of length l and constant linear mass density λ and hold it by its ends vertically, such that the rope takes on the form of two segments of length $\frac{l}{2}$ with a small[15] bend in the middle. Now, gently release one end while holding onto the other end. If the parts of the rope that cross the bend instantaneously come to a stop, determine $F(t)$, the force that you exert on the rope as a function of time.

31. *Pulling a Rope***

A long rope with linear mass density λ rests on a horizontal, frictionless table with a small[15] bend as shown in the figure on the next page. You grab

[15] Refer to footnote 14.

the end of the rope that is near the bend and begin to pull it. The length of the segment that initially crosses the bend is negligible and the entire rope is initially stationary (adapted from "Introduction to Classical Mechanics").



(1) Determine the force F required to pull the rope such that it moves at a constant velocity v.
(2) Determine the force, $F(t)$, required to pull the rope such that it moves at a constant acceleration a.
(3) Prove the reverse of part (1). If a constant force F is exerted on the rope, show that it travels at a constant velocity.
(4) If the force on the rope is now replaced with a spring such that the rightwards force on the rope is $k(L - x)$ where L is a constant and x is the rightwards displacement of the end that the spring is attached to, determine $x(t)$, the displacement of this end as a function of time for $x \leq L$.

Solutions

1. Regular Pentagon*

Draw two adjacent lines of symmetry of the pentagon. Their intersection is the center of mass. The interior angle of a pentagon is

$$\frac{3\pi}{5}.$$

By simple trigonometry, the perpendicular distance between the center of mass and the edge that a line of symmetry intersects with is

$$\frac{l}{2}\tan\frac{3\pi}{10}.$$

2. A Strange Rod*

Defining our origin to be at the left end of the rod,

$$\begin{aligned}\int xdm &= \int_0^l x\cdot\frac{k}{x+l}dx = \int_0^l k\left(1-\frac{l}{x+l}\right)dx \\ &= [kx - kl\ln|x+l|]_0^l = kl - kl\ln 2 \\ \int dm &= \int_0^l \frac{k}{x+l}dx = [k\ln|x+l|]_0^l = k\ln 2 \\ x_{CM} &= \frac{\int xdm}{\int dm} = l\left(\frac{1}{\ln 2}-1\right).\end{aligned}$$

3. Cone**

Let the mass density of the cone be ρ. The cross section between z and $z+dz$ is a disk of base area $\frac{z^2}{h^2}A$. Hence, the volume of this cross section is $\frac{z^2}{h^2}Adz$. Let the origin be located at the vertex of the cone and let the positive z-axis intersect the base of the cone perpendicularly. Then,

$$\begin{aligned}z_{CM} &= \frac{1}{M}\int_0^h zdm \\ &= \frac{1}{M}\int_0^h \rho\frac{z^3}{h^2}Adz\end{aligned}$$

$$= \frac{1}{M}\left[\frac{\rho z^4}{4h^2}A\right]_0^h$$
$$= \frac{Ah^2\rho}{4M}.$$

Furthermore,

$$M = \frac{1}{3}Ah\rho$$
$$\implies z_{CM} = \frac{3}{4}h.$$

Another way to solve this problem is to observe that we can form a conical layer of infinitesimal thickness by rotating an isosceles triangle of infinitesimal width and thickness, whose symmetrical axis subtends the half angle of the cone in question with the z-axis, for a complete revolution about the z-axis — implying that the center of mass of this conical layer lies at an altitude $\frac{l}{3}$ above its base where l is the conical layer's height. Now, we can retrieve a solid cone of height h by stacking various conical layers with heights ranging from $l = 0$ to $l = h$ on top of each other. Therefore, the center of mass of a solid cone is the weighted average of the center of masses of the different conical layers. The weight of each conical layer in this averaging process is proportional to its mass and hence its squared height l^2 (as mass is proportional to the surface area of the layer). Thus, the center of mass of a solid cone of height h lies at an altitude

$$\frac{\int_0^h l^2 \cdot \frac{l}{3} dl}{\int_0^h l^2 dl} = \frac{h}{4}$$

above its base, along its axis (due to symmetry).

4. Spherical Cap**

Define the origin to be at the center of the original sphere and let the positive z-axis pass through the vertex of the spherical cap. The center of mass of the cap must lie along the z-axis by symmetry. Now, we consider infinitesimal disks between coordinates z and $z + dz$. The radius of this disk r is given by the relationship

$$r^2 + z^2 = R^2$$
$$r^2 = R^2 - z^2.$$

Hence, the volume of this disk is $\pi r^2 dz = \pi(R^2 - z^2)dz$.

$$z_{CM} = \frac{1}{M}\int_{R-h}^{R} z dm = \frac{1}{M}\int_{R-h}^{R} \rho\pi(R^2 z - z^3)dz = \frac{\rho\pi h^2(2R-h)^2}{4M}.$$

The volume of the entire spherical cap is

$$V = \int_{R-h}^{R} \pi(R^2 - z^2)dz = \frac{\pi h^2(3R-h)}{3},$$
$$M = \rho V.$$

Therefore,

$$z_{CM} = \frac{3(2R-h)^2}{4(3R-h)}.$$

5. Triangle**

Let the plane of the triangle be the x'y'-plane. Define the origin O' at one vertex of the triangle and orientate the x'-axis such that another vertex lies on the x'-axis at coordinates $(x_2', 0)$. The last vertex has coordinates (x_3', y_3').

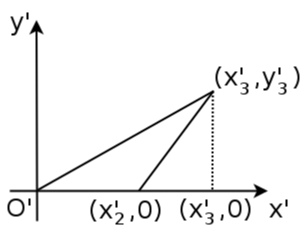


Figure 4.28: Triangle

The foot of the perpendicular of the third vertex on the x-axis is at coordinates $(x_3', 0)$ (it is external to the triangle in the figure above but its exact location does not matter). Then, we can determine the center of mass of the triangle via the center of masses of the two right-angled triangles formed — namely, one with vertices O', $(x_3', 0)$, (x_3', y_3') and another with vertices $(x_2', 0)$, $(x_3', 0)$ and (x_3', y_3'). In this case, we have to patch up the hole first to obtain the center of mass of the big triangle and then subtract the contribution due to the small triangle. Applying the result for the center of mass of a right-angled triangle, the coordinates of the center of masses are respectively $(\frac{2}{3}x_3', \frac{1}{3}y_3')$ and $(\frac{1}{3}x_2' + \frac{2}{3}x_3', \frac{1}{3}y_3')$. Their respective masses are $x_3' y_3'$ and $-(x_3' - x_2')y_3'$ (remember to include a negative sign for the second one as we are subtracting its contribution) where we have let the mass density of the triangle take a value of 2. Then, the coordinates of the center of mass of the general triangle are

$$\begin{pmatrix} x'_{CM} \\ y'_{CM} \end{pmatrix} = \frac{x'_3 y'_3 \begin{pmatrix} \frac{2}{3}x'_3 \\ \frac{1}{3}y'_3 \end{pmatrix} - (x'_3 - x'_2)y'_3 \begin{pmatrix} \frac{1}{3}x'_2 + \frac{2}{3}x'_3 \\ \frac{1}{3}y'_3 \end{pmatrix}}{x'_3 y'_3 - (x'_3 - x'_2)y'_3} = \begin{pmatrix} \frac{1}{3}x'_2 + \frac{1}{3}x'_3 \\ \frac{1}{3}y'_3 \end{pmatrix}.$$

In terms of the position vectors, $\boldsymbol{r}'_2$ and $\boldsymbol{r}'_3$, of the second and third vertices relative to the origin O',

$$\boldsymbol{r}'_{CM} = \frac{1}{3}\boldsymbol{r}'_2 + \frac{1}{3}\boldsymbol{r}'_3.$$

To determine the position vector of the center of mass $\boldsymbol{r}_{CM}$ with respect to a general origin O where the position vectors of the three vertices are $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$, observe that

$$\boldsymbol{r}_{CM} = \boldsymbol{r}'_{CM} + \boldsymbol{r}_1,$$

coupled with $\boldsymbol{r}'_2 = \boldsymbol{r}_2 - \boldsymbol{r}_1$ and $\boldsymbol{r}'_3 = \boldsymbol{r}_3 - \boldsymbol{r}_1$. Consequently,

$$\boldsymbol{r}_{CM} = \frac{1}{3}(\boldsymbol{r}_1 + \boldsymbol{r}_2 + \boldsymbol{r}_3).$$

For the second problem, observe that the midpoints of the edges of a triangle with vertices at $\boldsymbol{r}_1$, $\boldsymbol{r}_2$ and $\boldsymbol{r}_3$ have position vectors $\frac{\boldsymbol{r}_1+\boldsymbol{r}_2}{2}$, $\frac{\boldsymbol{r}_1+\boldsymbol{r}_3}{2}$ and $\frac{\boldsymbol{r}_2+\boldsymbol{r}_3}{2}$. Therefore, the center of mass of the new triangle formed by connecting the midpoints has position vector

$$\boldsymbol{r}_{CM,\,new} = \frac{1}{3}\left(\frac{\boldsymbol{r}_1 + \boldsymbol{r}_2}{2} + \frac{\boldsymbol{r}_1 + \boldsymbol{r}_3}{2} + \frac{\boldsymbol{r}_2 + \boldsymbol{r}_3}{2}\right) = \frac{1}{3}(\boldsymbol{r}_1 + \boldsymbol{r}_2 + \boldsymbol{r}_3) = \boldsymbol{r}_{CM,\,old},$$

which coincides with the center of mass of the original triangle. Therefore, the center of mass of the plated triangle is simply that of the first triangle and has coordinates $(\frac{x_1+x_2+x_3}{3}, \frac{y_1+y_2+y_3}{3})$.

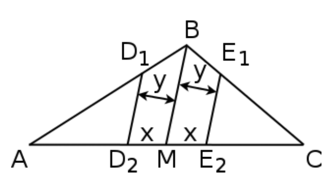


Figure 4.29: Triangle

To show that the center of mass of a triangle is the point of concurrency of the three medians, consider a particular median BM shown in Fig. 4.29. Consider two parallel lines, D_1D_2 and E_1E_2, that are equidistant from BM. Since $\triangle AD_1D_2 \sim \triangle ABM$, $\triangle CE_1E_2 \sim \triangle CBM$ and $\overline{AM} = \overline{MC}$ by the definition of a median, $\overline{D_1D_2} = \overline{E_1E_2}$. Therefore, the infinitesimal strips along D_1D_2 and E_1E_2 have equal masses. Furthermore, their perpendicular distances (y) to line BM are equal but they lie on opposite sides of line BM — implying that if we define the z-axis to be along line BM, the net contribution of these strips to the z-coordinate of the center of mass is zero. Repeating this argument for all parallel strips that are equidistant from BM, the z-coordinate of the center of mass of the entire triangle must be zero — signifying that it lies along line BM. Applying this procedure to the other two medians, we can prove that the center of mass must lie at the intersections of the medians.

6. Cylindrical Segment***

We will use Cartesian coordinates for our integrations, defining our origin to be at the left tip of the segment (side view). In Cartesian coordinates, dm is simply $\rho dxdydz$ where we define ρ to be the mass density of the wedge. We consider a infinitesimal element at coordinates (x, y, z).

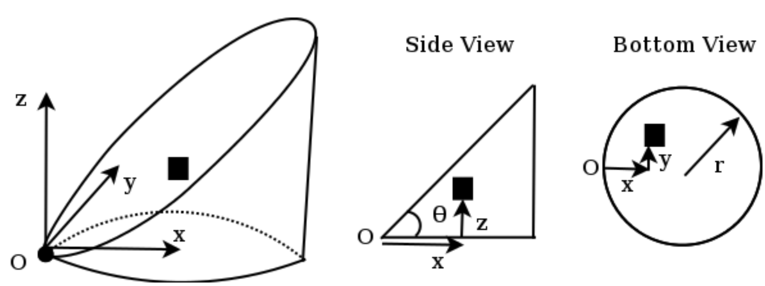


Figure 4.30: Cylindrical segment

However, we have to be extremely careful about our limits of integration for x, y and z. We see that for a given x, z can range from 0 to $x \tan\theta$ while y can range from $-\sqrt{r^2-(r-x)^2}$ to $\sqrt{r^2-(r-x)^2}$. The y-coordinate of the center of mass is trivially 0 due to symmetry. We shall calculate the other coordinates of the center of mass.

$$\begin{aligned}\int xdm &= \int_0^{2r}\int_{-\sqrt{r^2-(r-x)^2}}^{\sqrt{r^2-(r-x)^2}}\int_0^{x\tan\theta} x\rho dzdydx \\ &= \rho\int_0^{2r}\int_{-\sqrt{r^2-(r-x)^2}}^{\sqrt{r^2-(r-x)^2}} x^2\tan\theta dydx \\ &= \rho\tan\theta\int_0^{2r} 2x^2\sqrt{r^2-(r-x)^2}dx.\end{aligned}$$

The last integral can be evaluated using the substitution $r-x = r\sin\phi$, $dx = -r\cos\phi d\phi$.

$$\begin{aligned}&\int_0^{2r} x^2\sqrt{r^2-(r-x)^2}dx \\ &= -\int_{\frac{\pi}{2}}^{-\frac{\pi}{2}} (r^2-2r^2\sin\phi+r^2\sin^2\phi)r\cos\phi\cdot r\cos\phi d\phi \\ &= \frac{5}{8}\pi r^4.\end{aligned}$$

Thus,

$$\int xdm = \frac{5}{4}\rho\tan\theta\pi r^4.$$

Similarly we have

$$\begin{aligned}\int zdm &= \int_0^{2r}\int_{-\sqrt{r^2-(r-x)^2}}^{\sqrt{r^2-(r-x)^2}}\int_0^{x\tan\theta} z\rho dzdydx \\ &= \rho\int_0^{2r}\int_{-\sqrt{r^2-(r-x)^2}}^{\sqrt{r^2-(r-x)^2}} \frac{1}{2}x^2\tan^2\theta dydx \\ &= \rho\tan^2\theta\int_0^{2r} x^2\sqrt{r^2-(r-x)^2}dx \\ &= \frac{5}{8}\rho\tan^2\theta\pi r^4.\end{aligned}$$

Lastly, the volume of the cylindrical segment can be easily computed by observing that combining two of such segments forms a cylinder of radius r

and height $2r\tan\theta$. Hence, the total mass of this segment is

$$\begin{aligned} M &= \rho\pi r^3 \tan\theta, \\ x_{CM} &= \frac{\int x dm}{M} = \frac{5}{4}r, \\ y_{CM} &= 0, \\ z_{CM} &= \frac{\int z dm}{M} = \frac{5}{8}r\tan\theta. \end{aligned}$$

7. Constant Ratio***

The given property can be stated mathematically as

$$\frac{\int_0^y \lambda x dx}{y\int_0^y \lambda dx} = k.$$

Let $\int_0^y \lambda x dx = g(y)$ and $\int_0^y \lambda dx = m(y)$. Then,

$$g(y) = kym(y).$$

Differentiating the above equation with respect to y,

$$\lambda(y)y = km(y) + ky\lambda(y).$$

Rearranging,

$$\frac{\lambda(y)}{m(y)} = \frac{k}{(1-k)y}.$$

Since $\lambda(y) = \frac{dm}{dy}$, the left-hand side of the above equation is $\frac{d(\ln m)}{dy}$

$$\implies \int d(\ln m) = \int \frac{k}{(1-k)y}dy.$$

Integrating and simplifying,

$$m(y) = Ay^{\frac{k}{1-k}},$$

where A is a constant. A can be determined by substituting $m(l) = M$. Then,

$$\begin{aligned} A &= \frac{M}{l^{\frac{k}{1-k}}}, \\ m(y) &= \frac{M}{l^{\frac{k}{1-k}}}y^{\frac{k}{1-k}}, \\ \lambda(x) &= \frac{dm(x)}{dx} = \frac{kM}{(1-k)l^{\frac{k}{1-k}}}x^{\frac{2k-1}{1-k}}. \end{aligned}$$

A convenient limiting case to check for is $k = \frac{1}{2}$ as this simply corresponds to a stick of constant density. This is indeed the case as the exponent becomes $\frac{2k-1}{1-k} = \frac{2\cdot\frac{1}{2}-1}{1-\frac{1}{2}} = 0$.

8. Square Fractal***

Let the horizontal position of the center of mass be a distance $2x$ from the left edge of the largest square (the vertical position is trivial due to symmetry). Now, observe that the fractal is composed of a smaller fractal, two squares of length $\frac{l}{2}$ and the largest square of length l. The center of mass of the smaller fractal should be located a distance x from the left of its largest square by scaling arguments. Now, we simply need to determine the mass of this smaller fractal which is given by

$$\frac{l^2}{4} + \frac{3l^2}{16} + \frac{3l^2}{64} + \cdots = \frac{l^2}{2},$$

where we have let its mass density be one as its exact value is inconsequential. Now, the center of mass of the original fractal can be computed via that of the smaller fractal, two squares of length $\frac{l}{2}$ and the largest square.

$$\frac{\frac{l}{2}\cdot\frac{3l^2}{2} + (l+x)\cdot\frac{l^2}{2}}{\frac{3l^2}{2} + \frac{l^2}{2}} = 2x.$$

Solving,

$$2x = \frac{5}{7}l.$$

That said, there is an even more elegant method to computing the ratio of masses between the different components of this system. By scaling arguments, the original fractal should have four times the mass of the smaller fractal as the ratio of areas of corresponding squares is $4:1$. This means that the three remaining squares should have three times the mass of the smaller fractal. Then,

$$\frac{\frac{l}{2}\cdot 3 + (l+x)\cdot 1}{3+1} = 2x,$$

and we would again obtain $2x = \frac{5}{7}l$.

9. Triangle Fractal***

Let the vertical distance between the center of mass of the fractal and the bottom edge be $2y$. Observe that the original fractal is composed of two smaller fractals with halved length dimensions and a equilateral triangle of length $\frac{l}{2}$. By scaling arguments, the distance between the center of mass of the smaller fractals and their bottom edges should be y. Now, we just need to compute the ratio of relevant masses. By scaling arguments, the original fractal should have four times the mass of the smaller fractal. Therefore, the equilateral triangle of length $\frac{l}{2}$ should have two times the mass of the smaller fractal. Then,

$$\frac{y \cdot 1 + y \cdot 1 + \frac{\sqrt{3}}{6} l \cdot 2}{1+1+2} = 2y$$

$$2y = \frac{\sqrt{3}}{9} l.$$

10. Walking on a Plank*

There is no net external force on the system which comprises the person and the plank. Therefore, the center of mass of this combined system must not shift. Define the origin to be at the center of mass of the plank before the person begins moving. The coordinate of the person is $\frac{l}{2}$. The center of mass of the combined system is thus at $\frac{ml}{2(m+M)}$. Suppose that after the person moves to the other end of the plank, the center of mass of the plank is at coordinate x while that of the person is at $x - \frac{l}{2}$. Then,

$$\frac{m\left(x - \frac{l}{2}\right) + Mx}{m+M} = \frac{ml}{2(m+M)}$$

$$x = \frac{ml}{m+M}.$$

This can also be easily derived if we observe that the final state of the system is a horizontal flip of the initial state about the center of mass. Therefore, the displacement is twice the initial coordinate of the center of mass, $\frac{ml}{2(m+M)} \times 2 = \frac{ml}{m+M}$. The second problem is essentially the same as the first as an amount of mass m is transferred from one end of the plank to the other. Thus, the answer is still the same.

11. Pulling a Block*

The acceleration of the block is $a_b = \frac{F}{M}$. By Newton's third law, you also experience a force F and thus, acceleration $a_p = \frac{F}{m}$. These accelerations are opposite in direction. Therefore, in your frame, the block effectively accelerates at $a_b' = F(\frac{1}{M} + \frac{1}{m})$. Using the kinematics equation

$$l = \frac{1}{2}a_b' t^2,$$

$$t = \sqrt{\frac{2lMm}{F(M+m)}}.$$

12. Falling Slinky*

Define the origin to be at the top of the slinky and the y-axis to be positive downwards. For the first case, the bottom of the slinky is at $y = \frac{mg}{2k}$ while its center of mass is at $y = \frac{mg}{3k}$ (from previous results in this chapter). After the slinky is released, the only net external force on the slinky is its weight mg. Therefore, the center of mass of the slinky accelerates from rest at g. The bottom end only begins to move when the center of mass of the slinky reaches the bottom end, after traveling for a distance $\frac{mg}{2k} - \frac{mg}{3k} = \frac{mg}{6k}$ (i.e. the top end has accumulated the rest of the segments). The time required can be computed from basic kinematics.

$$\frac{1}{2}gt^2 = \frac{mg}{6k}$$

$$t = \sqrt{\frac{m}{3k}}.$$

In the second scenario, the entire slinky is stretched uniformly by an additional, overall length of $\frac{mg}{k}$ due to the additional weight attached. The coordinates of the bottom of the slinky and the center of mass are then $y = \frac{3mg}{2k}$ and $y = \frac{mg}{3k} + \frac{mg}{2k} = \frac{5mg}{6k}$ respectively. The time required is then

$$\frac{1}{2}gt^2 = \frac{3mg}{2k} - \frac{5mg}{6k} = \frac{2mg}{3k}$$

$$t = 2\sqrt{\frac{m}{3k}}.$$

13. Atwood's Machine 1*

In the case where the system is static, the forces on both m and M must be balanced. Letting the friction on M be f downslope (this is possibly a negative value),

$$T - mg = 0 \implies T = mg$$
$$T - Mg\sin\theta - f = 0$$
$$f = mg - Mg\sin\theta.$$

Since the normal force on M due to the slope is $N = Mg\cos\theta$, f must not exceed the maximum value of static friction given by

$$|f| \leq \mu_s N = \mu_s Mg\cos\theta$$
$$|mg - Mg\sin\theta| \leq \mu_s Mg\cos\theta.$$

This requires $M\sin\theta \leq m \leq M(\mu_s\cos\theta + \sin\theta)$ or $M(\sin\theta - \mu_s\cos\theta) \leq m \leq M\sin\theta$. If $m > M(\mu_s\cos\theta + \sin\theta)$ or $m < M(\sin\theta - \mu_s\cos\theta)$, the system can no longer remain static. In this case, let the acceleration of mass M in the direction parallel to the plane be a and the vertical acceleration of mass m be a_y, both taken to be positive upwards. Applying $F = ma$ to the blocks,

$$ma_y = T - mg$$
$$Ma = T - Mg\sin\theta - f$$
$$a = -a_y,$$

where the last equation is the conservation of string. Note that f could be $\mu_k N = \mu_k Mg\cos\theta$ or $-\mu_k N = -\mu_k Mg\cos\theta$, depending on the direction of relative motion between M and the slope. For now, we will stick with the variable f. Solving the above set of equations,

$$a_y = \frac{Mg\sin\theta - mg}{m+M} + \frac{f}{m+M},$$
$$a = \frac{mg - Mg\sin\theta}{m+m} - \frac{f}{m+M}.$$

To determine the sign of f, observe the sign of a in the case where there is no friction. If $m > M(\mu_s\cos\theta + \sin\theta)$ such that M moves and a without friction is positive, f tends to reduce the positive value of a and must hence undertake the positive value $\mu_k Mg\cos\theta$. Otherwise if $m < M(\sin\theta - \mu_s\cos\theta)$ such that M moves and a without friction is negative, f tends to increase the value of a and takes on the negative value $-\mu_k Mg\cos\theta$.

14. Atwood's Machine 2*

Let the tension in the string be T. For there to be no relative vertical motion, by considering $F = ma$ on mass m,

$$T - mg = 0 \implies T = mg.$$

Considering the three blocks as a whole system,

$$F = (M + \mu + m)a_x.$$

as the blocks must possess the same acceleration for there to be no relative motion. Lastly, by considering $F = ma$ on mass M,

$$T = Ma_x \implies a_x = \frac{mg}{M},$$

$$F = \frac{m(M + \mu + m)g}{M}.$$

15. Traveling Together**

Let the tension in the string be T, friction between the table and mass m_2 be f_1, the friction between m_2 and m_3 be f_2 and the common acceleration of the masses be a (we are using the conservation of string here without explicitly stating so). We obtain the following equations of motion while taking rightwards and downwards to be positive. Considering the system comprising m_1,

$$m_1 a = m_1 g - T.$$

From the system that contains m_2 and m_3,

$$(m_2 + m_3)a = T - f_1$$

$$\implies (m_1 + m_2 + m_3)a = m_1 g - f_1 = m_1 g - \mu_k(m_2 + m_3)g$$

$$a = \frac{m_1 - \mu_k(m_2 + m_3)}{m_1 + m_2 + m_3}g.$$

Next, by isolating m_3 alone,

$$|f_2| = m_3 a \leq \mu_s m_3 g$$

$$\implies \frac{m_1 - \mu_k(m_2 + m_3)}{m_1 + m_2 + m_3} \leq \mu_s.$$

The reason why we consider $|f_2| = m_3 a$ only and not $|f_2| = -m_3 a$ is due to the fact that $m_1 > \mu_0(m_2 + m_3) > \mu_k(m_2 + m_3)$ for the set-up to move in the first place, where μ_0 is the coefficient of static friction between the table and m_2. Solving,

$$m_1 \leq \frac{(\mu_s + \mu_k)(m_2 + m_3)}{1 - \mu_s}.$$

16. Sliding down a Plane**

Define rightwards and upwards to be the positive directions. Let N denote the normal force on the block due to the plane (this is not necessarily $mg\cos\theta$), a_x and a_y be the horizontal and vertical accelerations of the block and A be the horizontal acceleration of the plane. Writing the $F = ma$ equations in the vertical and horizontal directions,

$$N\cos\theta - mg = ma_y,$$
$$N\sin\theta = ma_x,$$
$$-N\sin\theta = MA.$$

For the block to stay on the plane,

$$\frac{a_y}{a_x - A} = -\tan\theta.$$

Solving the above equations yields

$$N = \frac{g}{\sin\theta\tan\theta(\frac{1}{m} + \frac{1}{M}) + \frac{\cos\theta}{m}},$$
$$A = -\frac{N\sin\theta}{M} = -\frac{mg\sin\theta\cos\theta}{M + m\sin^2\theta}.$$

17. Atwood's Machine 3**

If we let the tension in the string holding m_1 be T, the tension in the string holding m_2 is $4T$. Thus, we can obtain the equations of motion

$$m_1 a_1 = m_1 g - T,$$
$$m_2 a_2 = m_2 g - 4T,$$

where we have chosen downwards to be positive. To determine the conservation of string equation, imagine if the second and fourth pulleys on top (counting from the left) were displaced downwards by a and b respectively. If we let the displacements of m_1 and m_2 be x_1 and x_2 respectively, we can observe that

$$x_1 + 2a + 2b = 0$$

as the second and fourth pulleys release an amount of string that is twice their displaced distances. Furthermore, by considering the segment of string holding up the pulley that is connected to m_2, we obtain

$$a + b = 2x_2,$$

as moving m_2 downwards by x_2 requires additional string of $2x_2$ in length.

$$4x_2 + x_1 = 0 \implies a_1 + 4a_2 = 0.$$

Solving,

$$T = \frac{5m_1m_2g}{16m_1 + m_2},$$

$$a_1 = \frac{16m_1 - 4m_2}{16m_1 + m_2}g,$$

$$a_2 = \frac{m_2 - 4m_1}{16m_1 + m_2}g.$$

18. Atwood's Machine 4*

The magnitude of tension is constant throughout the string and we will define it to be T. Writing the $F = ma$ equation for both masses and taking downwards to be positive,

$$ma_1 = mg - T,$$

$$2ma_2 = 2mg - 4T.$$

Lastly, we derive our conservation of string equation. If the right mass moves by a distance d, the left mass has to move a distance $4d$ as there are four string segments connected to the right mass as compared to the one on the left mass. Thus,

$$a_1 + 4a_2 = 0.$$

We solve for T by adding the first equation to twice of the second.

$$m(a_1 + 4a_2) = 5mg - 9T \implies T = \frac{5}{9}mg,$$

$$a_1 = \frac{4}{9}g,$$

$$a_2 = -\frac{1}{9}g.$$

19. Atwood's Machine 5**

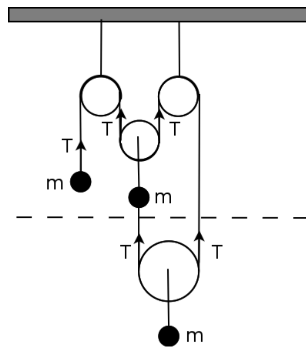


Figure 4.31: Atwood's machine 4

Writing our $F = ma$ equations,

$$
\begin{aligned}
ma_1 &= mg - T, \\
ma_2 &= mg - T, \\
ma_3 &= mg - 2T,
\end{aligned}
$$

where the masses are numbered from left to right. To derive our conservation of string equation, assume that the middle mass moves above the dotted line by a distance d. Then, string of length $2d$ will be made available elsewhere. However, d length of string is required to be "added" to the string connecting the middle mass and the bottom pulley (above the dotted line). Thus, only an additional d length of string remains to be shared between the pulleys connected to the other masses. We let x and y be the lengths of string that the pulleys of the left and right masses gain, respectively. Then, the right mass only moves down by $\frac{y}{2}$ as there must be an increase in the length of the string on each side of the pulley by $\frac{y}{2}$. We can then express the accelerations of our masses as the second time derivatives of x, y and d.

$$
\begin{aligned}
x + y &= d, \\
a_2 &= -\ddot{d}, \\
a_1 &= \ddot{x}, \\
a_3 &= \frac{\ddot{y}}{2}, \\
\implies a_1 + a_2 + 2a_3 &= 0.
\end{aligned}
$$

Solving,

$$T = \frac{2}{3}mg,$$

$$a_1 = \frac{1}{3}g,$$

$$a_2 = \frac{1}{3}g,$$

$$a_3 = -\frac{1}{3}g.$$

20. Atwood's Machine 6**

Number the masses from left to right in ascending order. The key observation in this problem is that the tension in the string above the middle mass is not necessarily equal to the tension in the string below it as they are disjoint. If we let the tension on top of the middle mass be T_1 and that below it be T_2,

$$ma_1 = mg - T_1,$$

$$ma_2 = mg + T_2 - T_1,$$

$$ma_3 = mg - 2T_2.$$

In this case, we actually have two conservation of string equations as there are two string segments.

$$a_2 = -a_1,$$

$$2a_3 = a_2.$$

Solving,

$$T_1 = \frac{11}{9}mg,$$

$$T_2 = \frac{4}{9}mg,$$

$$a_1 = -\frac{2}{9}g,$$

$$a_2 = \frac{2}{9}g,$$

$$a_3 = \frac{1}{9}g.$$

21. Pulling a Mass**

Since pivot A is fixed, the mass can only have a velocity tangential to the first string (let this tangential speed be u anti-clockwise).

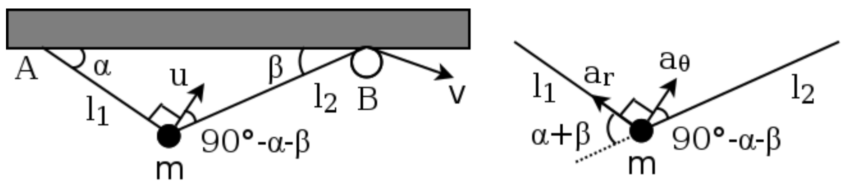


Figure 4.32: Velocity and acceleration of mass

The component of the ball's velocity along the second string contributes to the rate of decrease of the length of segment connecting the mass to the fixed pulley B, $-\dot{l}_2 = v$ (consider polar coordinates about the fixed pulley B). Thus,

$$u\cos(90^\circ - \alpha - \beta) = v$$
$$u = \frac{v}{\sin(\alpha+\beta)}.$$

Next, let the radial and tangential accelerations of the mass relative to the first string be a_r and a_θ as labelled in Figure 4.32. Since pivot A is fixed and the length of the first string must be maintained, a_r must correspond to the centripetal acceleration.

$$a_r = \frac{u^2}{l_1} = \frac{v^2}{l_1 \sin^2(\alpha+\beta)}.$$

Similarly, since $\ddot{l}_2 = 0$ in the equation for radial acceleration in polar coordinates about the fixed pulley B, the component of acceleration parallel to the second string must also provide the centripetal acceleration associated with the velocity $u\cos(\alpha+\beta)$ tangential to the second string.

$$a_\theta \sin(\alpha+\beta) - a_r\cos(\alpha+\beta) = \frac{u^2\cos^2(\alpha+\beta)}{l_2}.$$

Substituting the expressions for u and a_r,

$$a_\theta = \frac{v^2\cos(\alpha+\beta)}{l_2\sin^3(\alpha+\beta)}\left(\frac{1}{l_1} + \frac{\cos(\alpha+\beta)}{l_2}\right).$$

Finally, we consider the forces on the mass. Let the tension on the mass due to the second string be T_2 — the tension due to the other string is not of concern. Considering the components of T_2 and the weight of the mass perpendicular to the first string,

$$T_2 \sin(\alpha + \beta) - mg \cos \alpha = ma_\theta$$

$$\implies T_2 = \frac{mg \cos \alpha}{\sin(\alpha + \beta)} + \frac{mv^2 \cos(\alpha + \beta)}{l_2 \sin^4(\alpha + \beta)} \left(\frac{1}{l_1} + \frac{\cos(\alpha + \beta)}{l_2} \right).$$

22. Equivalent Mass**

The accelerations of the pulley and the equivalent mass must be equal so that the string holding them will be conserved for all external set-ups connected to them. Let this common acceleration be a_{eq}. Next, for the set-ups to be completely equivalent, the tensions in the strings holding the pulley and the equivalent mass must be identical. Let this common tension be T. Next, let the accelerations of m_1 and m_2 be $a_{eq} + a'$ and $a_{eq} - a'$ (note that their average must be a_{eq}). Then,

$$m_1 g - \frac{T}{2} = m_1(a_{eq} + a'),$$
$$m_2 g - \frac{T}{2} = m_2(a_{eq} - a'),$$
$$m_{eq} g - T = m_{eq} a_{eq}.$$

Solving,

$$m_{eq} = \frac{4m_1 m_2}{m_1 + m_2},$$
$$T = \frac{4m_1 m_2 (g - a_{eq})}{m_1 + m_2},$$
$$a' = \frac{m_1 - m_2}{m_1 + m_2}(g - a_{eq}).$$

Note that a_{eq}, and thus T and a', is indeterminate as it depends on the set-up that these masses are connected to.

23. Infinite Atwood's Machine***

(a) Method 1: Equivalent Masses

We have shown in the previous problem that the equivalent mass is given by

$$m_{eq} = \frac{4m_1 m_2}{m_1 + m_2}.$$

We can apply this result in a bottom-up fashion such that m_1 is maintained at m while m_2 is the current equivalent mass as we proceed bottom-up. Since m_2 begins at m, we are equivalently applying the following function to m infinitely many times (in a layered fashion).

$$f(x) = \frac{4mx}{m+x}.$$

Notice that $f(x) > x$ as long as $x < 3m$ and that $f(3m) = 3m$ is a stationary point. Since we are beginning at m and each iteration increases the value of the equivalent mass until $3m$, the entire connection on the right of the top pulley is equivalent to a mass $3m$ — reducing the problem to the simplest Atwood's machine with two masses that we have already solved for. If we start from 1 and number the masses from left to right, directly applying the result of the first problem in Sec. 4.6.3 would yield

$$a_1 = -\frac{g}{2}$$

where the positive direction is downwards. Let the tension on the first mass be T. Then, the tension on the kth mass would be $\frac{T}{2^{k-1}}$. Applying $F = ma$ to the first and kth mass,

$$mg - T = ma_1 \implies T = \frac{3mg}{2}$$

$$mg - \frac{T}{2^{k-1}} = ma_k$$

$$a_k = \left(1 - \frac{3}{2^k}\right) g.$$

(b) Method 2: Scaling Arguments and Symmetry

Let the tension on the first mass be T again. Then the tension in the string connecting the first pulley to the wall is $2T$ while the tension in the string connecting the second pulley to the first mass is T. We note that if our system of masses is fixed, the tension between the first pulley and the wall, $2T$, can only depend on the gravitational field strength g and should in fact be proportional to g, by dimensional analysis. Thus,

$$\frac{2T}{g} = k$$

for some constant k as long as we keep the same system of masses, though the gravitational field strength may vary. Now let us consider the system of masses on the right end of the first pulley. The second pulley is accelerating at $-a_1$, with a_1 taken to be positive downwards. Thus if we consider the frame of the second pulley, it is as if the system of masses below it existed in a world where gravity is $g + a_1$. Furthermore, the tension in the string holding the second pulley is akin to that in the string connecting the first pulley to the wall as the second pulley does not know what it is connected to. Thus, by leveraging this symmetry and the fact that the tension should only depend on the apparent gravitational field strength,

$$\frac{2T}{g} = \frac{T}{g + a_1}.$$

We can then easily obtain the answer

$$a_1 = -\frac{g}{2}.$$

Considering the $F = ma$ equations of the first and kth mass would again lead to the result

$$a_k = \left(1 - \frac{3}{2^k}\right) g.$$

Actually, $T = 0$ is also a solution to the above equation, but that would mean that all masses would be in free fall. This happens if we choose the last mass to be a massless pulley as this implies that the tension in the last string is zero which causes the tension in every string to be zero. Method one will also result in the same conclusion as the equivalent mass will obviously be zero. Therefore, a single mass can surprisingly have such a large impact!

24. Roller Coaster*

For the roller coaster to lose contact with the track, the normal force on the roller coaster due to the track must become zero. Now, observe that for an angular position θ clockwise from the top of the circle, the net external force on the roller coaster in the radial direction is

$$F_r = -mg\cos\theta - N,$$

which must be equal to the required centripetal force $-\frac{mv^2}{R}$. Let l be the vertical height that the roller coaster has fallen. By the conservation of energy, $2gl = v^2$. Therefore, at the top of the circle where $\theta = 0$, the required centripetal force is the smallest and yet the component of its weight in the radial direction is the largest — meaning that the normal force at this juncture must be the smallest. Thus, the roller coaster has the greatest tendency to lose contact at the top of the circle. At this point,

$$F_r = -mg - N$$
$$v^2 = 2g(H - 2R)$$
$$\implies mg + N = \frac{2mg(H-2R)}{R}.$$

For the roller coaster to not lose contact with the surface,

$$N = \frac{2mg(H-2R)}{R} - mg > 0$$
$$\implies H > \frac{5}{2}R.$$

25. Falling off a Circle**

We set our origin at the center of the circle and define θ to be the angular coordinate of the particle from the vertical, while taking the clockwise direction to be positive. The net force in the radial direction must provide the required centripetal force.

$$N - mg\cos\theta = -\frac{mv^2}{R}.$$

From the conservation of energy,

$$\frac{v^2}{2} = gR(1 - \cos\theta).$$

At the instance where the particle just begins to slip, $N = 0$. Substituting this condition and the equation above into the first equation,

$$mg\cos\theta = 2mg - 2mg\cos\theta_c \implies \theta_c = \cos^{-1}\left(\frac{2}{3}\right).$$

26. Stationary Stand**

Let θ be the instantaneous angular coordinate of m from the vertical, defined to be positive anti-clockwise. Let v and T denote the speed of m and tension in the string at angle θ. By the conservation of energy,

$$mgl\cos\theta = \frac{1}{2}mv^2.$$

The net force on m in the radial direction is

$$F_r = mg\cos\theta - T.$$

This must be equal to the required centripetal force $-\frac{mv^2}{l} = -2mg\cos\theta$. Thus,

$$T = 3mg\cos\theta.$$

Let N be the normal force on the stand due to the table and f be the friction between the stand and the table. For the forces on the stand to be balanced,

$$N = Mg + T\cos\theta = Mg + 3mg\cos^2\theta,$$
$$f = T\sin\theta = 3mg\sin\theta\cos\theta.$$

Lastly, we have the condition

$$|f| \leq \mu N = \mu(Mg + 3mg\cos^2\theta)$$
$$\mu \geq \frac{|3m\sin\theta\cos\theta|}{M + 3m\cos^2\theta},$$

where $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. We can discard the absolute value and determine the maxima of the right-hand side for $\theta \geq 0$ by differentiation as the situation is symmetrical. Doing so will yield the maxima condition to be

$$\tan\theta = \sqrt{1 + \frac{3m}{M}}.$$

At this value of θ,

$$\mu \geq \frac{3m\sqrt{1 + \frac{3m}{M}}}{2M + 6m}.$$

27. Rotating Rod**

Consider an infinitesimal segment between r and $r+dr$. Denote the outwards direction to be positive. T, the tension at r, points radially inwards while at $r + dr$, it points radially outwards and is denoted by $T + dT$. The net force in the radial direction must provide the centripetal force.

$$dT = -\rho r\omega^2 dr$$
$$\implies T(r) = -\frac{\rho r^2\omega^2}{2} + c,$$

where c is determined by a boundary condition. The tension at free ends must be zero. Hence, if $r = 0$ is a free end, $c = 0$.

$$T(r) = -\frac{\rho r^2\omega^2}{2}.$$

If $r = l$ is a free end, $c = \frac{\rho l^2\omega^2}{2}$.

$$T(r) = \frac{\rho\omega^2}{2}(l^2 - r^2).$$

28. Rotating Chain**

Consider an infinitesimal segment of a chain of length $Rd\theta$ between θ and $\theta + d\theta$, in the plane containing the chain in polar coordinates. Figure 4.33 depicts the free-body diagram of this segment.

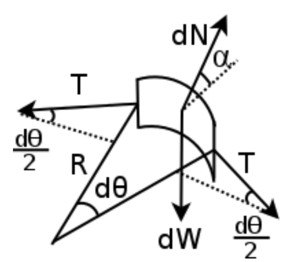


Figure 4.33: Free-body diagram of an infinitesimal segment

T represents the tension at the two ends of this segment due to neighboring segments (they must be equal for the segment to not accelerate tangentially). dN is the normal force on the segment due to the cone, which is directed at the half angle α above the plane containing the chain due to the inclination of the cone. dW is the weight of this segment, which is equal to $\lambda Rgd\theta$. For the forces to be balanced in the vertical direction,

$$dN \sin\alpha = dW$$
$$\implies dN = \frac{\lambda Rg}{\sin\alpha} d\theta.$$

The radial component of the tensions is $2T\sin\frac{d\theta}{2} \approx Td\theta$ using the small angle approximation $\sin x \approx x$ for small x. The net force in the radial direction is then

$$F_r = dN\cos\alpha - Td\theta = (\lambda Rg\cot\alpha - T)d\theta.$$

This must be equal to the required centripetal force, so

$$\begin{aligned}(\lambda Rg\cot\alpha - T)d\theta &= -\lambda Rd\theta R\omega^2\\ T &= \lambda R^2\omega^2 + \lambda Rg\cot\alpha.\end{aligned}$$

29. Sweeping Pan**

The mass of the dustpan after it has travelled a distance x is

$$m(x) = M + \sigma lx.$$

As the pan travels with a constant acceleration a,

$$\begin{aligned}x &= \frac{1}{2}at^2,\\ v &= at,\\ m &= M + \frac{1}{2}\sigma lat^2.\end{aligned}$$

The momentum of the pan at a particular time t is

$$\begin{aligned}p = mv &= Mat + \frac{1}{2}\sigma la^2t^3,\\ F = \frac{dp}{dt} &= Ma + \frac{3}{2}\sigma la^2t^2.\end{aligned}$$

30. Holding a Rope**

The important observation is that there must be no tension in the rope as the small bend is essentially horizontal and any tension in the rope would cause it to accelerate with infinite magnitude upwards (tension on both sides is directed upwards). Therefore, the moving part of the rope is under free fall.

After time t, the free end of the rope would have dropped by a vertical distance $\frac{gt^2}{2}$. However, only $\frac{gt^2}{4}$ amount of rope crosses over the bend (similar to a movable pulley). Then, a remaining $\frac{l}{2} - \frac{gt^2}{4}$ length of rope travels with velocity gt, with the positive direction defined to be downwards. Thus, the momentum of the entire rope is

$$p = \lambda \left(\frac{l}{2} - \frac{gt^2}{4} \right) gt.$$

The net force on the entire rope, which comprises its weight and the force exerted by you, must be equal to its rate of change of momentum.

$$\lambda gl - F = \frac{dp}{dt} = \frac{\lambda gl}{2} - \frac{3\lambda g^2 t^2}{4}$$

$$F = \frac{\lambda gl}{2} + \frac{3\lambda g^2 t^2}{4}.$$

Note that at $t = \sqrt{\frac{2l}{g}}$, F drops abruptly from $2\lambda gl$ to λgl (the weight of the rope) as there is no longer any need to provide an upwards acceleration to parts of the rope that are initially traveling downwards and cross over the bend, to ensure that they are stationary at the other side of the bend.

31. Pulling a Rope**

At any point in time, only rope segments that have crossed the bend are moving as the rope is slack everywhere (refer to the explanation in the previous answer). After pulling the rope for a distance x, only $\frac{x}{2}$ amount of rope will be traveling at $\dot{x}$ (similar to the case of a movable pulley). Thus, the momentum of the entire rope is

$$p(x) = \frac{\lambda x \dot{x}}{2}.$$

The relationship between the net force on the rope and its rate of change of momentum is

$$F = \frac{dp}{dt} = \frac{\lambda \dot{x}^2}{2} + \frac{\lambda x \ddot{x}}{2}.$$

For the first scenario, $\dot{x} = v$ and $\ddot{x} = 0$. The required force is

$$F = \frac{\lambda v^2}{2}.$$

In the second scenario, $x = \frac{1}{2}at^2$, $\dot{x} = at$ and $\ddot{x} = a$. The required force is

$$F = \frac{3\lambda a^2 t^2}{4}.$$

In the third scenario, F is a constant which implies that $p = Ft$.

$$\frac{\lambda x}{2}\frac{dx}{dt} = Ft$$

$$\int_0^x \frac{\lambda}{2} x dx = \int_0^t F t dt$$

$$\frac{\lambda x^2}{4} = \frac{Ft^2}{2}$$

$$x = \sqrt{\frac{2F}{\lambda}} t$$

$$\dot{x} = \sqrt{\frac{2F}{\lambda}}.$$

In the last scenario, $F = k(L - x)$.

$$k(L-x) = \frac{\lambda \dot{x}^2}{2} + \frac{\lambda x}{4}\frac{d\dot{x}^2}{dx},$$

where we have used the trick $\ddot{x} = \frac{1}{2}\frac{d\dot{x}^2}{dx}$. Multiplying the above by the appropriate integrating factor $4x$,

$$2\lambda x \dot{x}^2 + \lambda x^2 \frac{d\dot{x}^2}{dx} = \lambda \frac{d(x^2\dot{x}^2)}{dx} = 4k(L-x)x.$$

Integrating and substituting the limits $x = 0$ and $\dot{x} = 0$ when $t = 0$,

$$\lambda \int_0^{x^2\dot{x}^2} d(x^2\dot{x}^2) = \int_0^x 4k(L-x)x dx$$

$$\lambda x^2 \dot{x}^2 = 2kLx^2 - \frac{4kx^3}{3}$$

$$\dot{x} = \sqrt{\frac{2kL}{\lambda} - \frac{4kx}{3\lambda}}.$$

The positive square root is taken as $\dot{x} > 0$ for $0 < x \leq L$ as $\ddot{x} \geq 0$.

$$\int_0^x \frac{1}{\sqrt{\frac{3}{2}L - x}} dx = \int_0^t \sqrt{\frac{4k}{3\lambda}} dt$$

$$\sqrt{6L} - \sqrt{6L - 4x} = \sqrt{\frac{4k}{3\lambda}} t$$

$$6L - 4x = \left(\sqrt{6L} - \sqrt{\frac{4k}{3\lambda}}t\right)^2$$

$$x = \frac{3}{2}L - \left(\sqrt{\frac{3L}{2}} - \sqrt{\frac{k}{3\lambda}}t\right)^2.$$

This is valid until $x = L$, after which the force on the rope becomes compressive — causing the rope to deform.

Chapter 5

Rotational Dynamics

In this chapter, we will introduce concepts such as angular momentum and torques, to analyze the fixed axis rotation of rigid bodies.

5.1 Angular Momentum and Torque

The angular momentum of a particle $\boldsymbol{L}$ with respect to a certain origin is defined as:

$$\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}, \tag{5.1}$$

where $\boldsymbol{r}$ and $\boldsymbol{p}$ refer to the position vector and momentum of the particle respectively. The total angular momentum of a system of particles is simply the sum of the individual angular momenta. Before we understand how the angular momentum is a rotational attribute of a system of particles, let us first calculate the angular momentum of a rigid body.

5.1.1 *Rigid Body about Stationary Axis*

We shall now derive the expression for the angular momentum of a rigid object with a continuous mass distribution, with an angular velocity $\boldsymbol{\omega}$ in the z-direction. The positive direction of the angular velocity vector, $\boldsymbol{\omega}$, is determined by the right-hand corkscrew rule (point your right thumb in the positive z-direction, your other fingers will curl in the positive direction of $\boldsymbol{\omega}$).

In the case where we are able to identify an ICOR,[1] we can define the origin at the ICoR. Then, the z-component of the angular momentum of the

[1]Since an instantaneous axis of rotation exists in three-dimensional set-ups, we shall refer to a point on the instantaneous axis as an ICoR too. Remember that both the instantaneous axis of rotation and the ICoR can be external to the rigid body. Refer to Sec. 3.5 for revision.

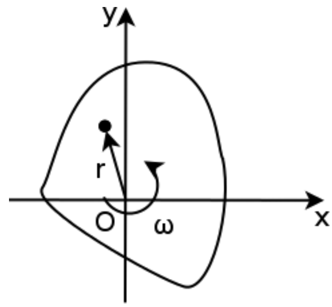


Figure 5.1: Rigid body rotating about the z-axis

entire rigid body takes on a rather convenient form:

$$L_z = I\omega, \tag{5.2}$$

where

$$I = \int (x^2 + y^2) dm = \int r_\perp^2 dm \tag{5.3}$$

is known as the moment of inertia about the z-axis and $r_\perp$ is the perpendicular distance between an infinitesimal mass element and the z-axis. Furthermore, the total kinetic energy of the rigid body is

$$T = \frac{1}{2} I\omega^2. \tag{5.4}$$

As the angular velocity $\boldsymbol{\omega}$ does not change if we switch between frames with relatively non-rotating axes, we can apply the above equations by switching to a frame co-moving with a particle on the rigid body and defining the origin at the particle such that it is an ICoR. We shall refer to this process as fixing an origin to the rigid body, but keep in mind that the axes of the co-moving frame must not be rotating in the original frame (as that would change the observed angular velocity). Furthermore, the angular momentum will be computed with respect to this co-moving frame and not the original frame. Finally, a major detriment of such an approach is that the co-moving frame is most probably accelerating, as the velocity of the particle that it follows may change — creating future complications (e.g. in relating torque and the rate of change of angular momentum).

Proof: Recall that the velocity of a point on a rigid body is given by

$$\boldsymbol{v} = \boldsymbol{v}_{ref} + \boldsymbol{\omega} \times \boldsymbol{r},$$

where $\boldsymbol{v}_{ref}$ is the velocity of a reference point on the rigid body and $\boldsymbol{r}$ is the vector pointing from this reference point to the point of concern. If we define the reference point to be an ICoR whose $\boldsymbol{v}_{ref} = \mathbf{0}$ by definition, the velocity of a point on a rigid body is

$$\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}.$$

Let us first consider the case where the particles on the rigid body are discrete before moving onto the continuous case. If there are N particles on the rigid body in total, the total angular momentum is determined by the sum of the angular momenta of each individual particle.

$$\begin{aligned}
\boldsymbol{L} &= \sum_{i=1}^{N} \boldsymbol{r}_i \times m_i \boldsymbol{v}_i \\
&= \sum_{i=1}^{N} \boldsymbol{r}_i \times m_i (\boldsymbol{\omega} \times \boldsymbol{r}_i) \\
&= \sum_{i=1}^{N} \boldsymbol{r}_i \times m_i \begin{pmatrix} 0 \\ 0 \\ \omega \end{pmatrix} \times \begin{pmatrix} r_{ix} \\ r_{iy} \\ r_{iz} \end{pmatrix} \\
&= \sum_{i=1}^{n} m_i \begin{pmatrix} r_{ix} \\ r_{iy} \\ r_{iz} \end{pmatrix} \times \begin{pmatrix} -r_{iy}\omega \\ r_{ix}\omega \\ 0 \end{pmatrix} \\
&= \sum_{i=1}^{n} m_i \begin{pmatrix} -r_{ix} r_{iz} \\ -r_{iy} r_{iz} \\ (r_{ix}^2 + r_{iy}^2) \end{pmatrix} \omega.
\end{aligned}$$

Notice that the angular momentum vector does not generally point in the direction of the angular velocity vector.[2] However, if we just consider the z-component[3] of $\boldsymbol{L}$,

$$L_z = \sum_{i=1}^{n} m_i r_{i\perp}^2 \omega,$$

[2] However, they are generally aligned if the object is flat and lies in the xy-plane such that r_{iz} is zero for all i.

[3] If we are only concerned with the z-component in the first place, a quicker way of deriving this is by applying the "BAC-CAB" rule (Eq. (3.6)) such that $\boldsymbol{L} = \int \boldsymbol{r} \times (\boldsymbol{\omega} \times \boldsymbol{r}) dm = \int [\boldsymbol{\omega}(\boldsymbol{r} \cdot \boldsymbol{r}) - \boldsymbol{r}(\boldsymbol{\omega} \cdot \boldsymbol{r})]\, dm$. The z-component of the second term in the integrand is simply $-r_z^2 \omega$ such that the z-component of the integral becomes $L_z = \int (r^2 - r_z^2)\omega dm = \int r_\perp^2 \omega dm$.

where $r_{i\perp}$ is the perpendicular distance between the ith particle and the z-axis. In the limit where $n \to \infty$ and $m_i \to 0$,

$$L_z = \int r_\perp^2 \, dm \, \omega.$$

The integral is evaluated over the entire mass distribution. The term $\int r_\perp^2 \, dm$ is known as the moment of inertia I, with respect to the z-axis. As its nomenclature implies, it is a measure of its rotational inertia — as we shall see in a later section.

$$I = \int r_\perp^2 \, dm.$$

Thus,

$$L_z = I\omega.$$

Let us derive the total kinetic energy of the object. We shall use the fact that the kinetic energy of a particle with mass m and velocity v is $\frac{1}{2}mv^2$, though we have not formally introduced it yet. Then, the total kinetic energy T is

$$\begin{aligned}
T &= \int \frac{1}{2}\boldsymbol{v} \cdot \boldsymbol{v} dm \\
&= \frac{1}{2}\int \begin{pmatrix} -r_y\omega \\ r_x\omega \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -r_y\omega \\ r_x\omega \\ 0 \end{pmatrix} dm \\
&= \frac{1}{2}\int (r_x^2 + r_y^2) dm \, \omega^2 \\
&= \frac{1}{2}\int r_\perp^2 \, dm \, \omega^2 \\
&= \frac{1}{2} I\omega^2.
\end{aligned}$$

In conclusion, if we are able to identify an ICoR in the current frame (such as a point at which the rigid object is pivoted about), Eqs. (5.2) and (5.4) can be used to determine the angular momentum and total kinetic energy of a rigid body. Note that even though a point on the rigid body may be moving in the lab frame, we can still apply the above results in its co-moving frame ($\boldsymbol{\omega}$ remains the same) by defining the origin at that point, such that it is an ICoR. Thus, the equations above are applicable with respect to all origins fixed to the body.

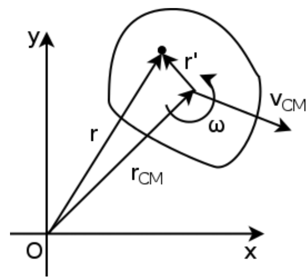


Figure 5.2: Rigid body

5.1.2 *Rigid Body about General Axis*

In the more general case, our origin can be a stationary point in the current frame that is not an ICoR. Then, the z-component of the angular momentum of the body is

$$L_z = I_{CM}\omega + M\left(\boldsymbol{r}_{CM} \times \boldsymbol{v}_{CM}\right)_z, \tag{5.5}$$

where I_{CM} is the moment of inertia of the rigid body about an axis passing through its center of mass, parallel to the z-axis and M is the total mass of the rigid body. $\boldsymbol{r}_{CM}$ and $\boldsymbol{v}_{CM}$ are the position vector and velocity of the center of mass, respectively. The total kinetic energy of the rigid body is

$$T = \frac{1}{2}I_{CM}\omega^2 + \frac{1}{2}Mv_{CM}^2. \tag{5.6}$$

Proof: Now that the origin is not an ICoR, we cannot write $\boldsymbol{v} = \boldsymbol{\omega} \times \boldsymbol{r}$. However, we can still use $\boldsymbol{v} = \boldsymbol{v}_{ref} + \boldsymbol{\omega} \times \boldsymbol{r}$. The convenient reference point to choose, in this case, is the center of mass, as we shall soon see. Define $\boldsymbol{r}_{CM}$ to be the position vector of the center of mass and $\boldsymbol{r}'$ to be the vector pointing from the center of mass to the point of concern (Fig. 5.2). Then, the velocity of a point on the rigid body can be expressed as

$$\boldsymbol{v} = \boldsymbol{v}_{CM} + \boldsymbol{\omega} \times \boldsymbol{r}'.$$

Its position vector can similarly be expressed as

$$\boldsymbol{r} = \boldsymbol{r}_{CM} + \boldsymbol{r}'.$$

From the definition of angular momentum,

$$\begin{aligned}\boldsymbol{L} &= \int (\boldsymbol{r}_{CM} + \boldsymbol{r}') \times (\boldsymbol{v}_{CM} + \boldsymbol{\omega} \times \boldsymbol{r}')dm \\ &= \int \boldsymbol{r}_{CM} \times \boldsymbol{v}_{CM} dm + \boldsymbol{r}_{CM} \times \left(\int \boldsymbol{\omega} \times \boldsymbol{r}' dm\right)\end{aligned}$$

$$+\left(\int \boldsymbol{r}'dm\right)\times\boldsymbol{v}_{CM}+\int \boldsymbol{r}'\times\left(\boldsymbol{\omega}\times\boldsymbol{r}'\right)dm$$

$$=M\boldsymbol{r}_{CM}\times\boldsymbol{v}_{CM}+\boldsymbol{r}_{CM}\times\left(\boldsymbol{\omega}\times\int\boldsymbol{r}'dm\right)$$

$$+\left(\int \boldsymbol{r}'dm\right)\times\boldsymbol{v}_{CM}+\int \boldsymbol{r}'\times\left(\boldsymbol{\omega}\times\boldsymbol{r}'\right)dm,$$

where M is the total mass of the rigid body. The second and third terms are equal to zero by the definition of the center of mass as $\int \boldsymbol{r}'dm = \int(\boldsymbol{r}-\boldsymbol{r}_{CM})dm = M\boldsymbol{r}_{CM} - M\boldsymbol{r}_{CM} = 0$. Furthermore, by comparing the last term to the expression for the angular momentum of a rigid body about a stationary axis in the previous section, it can be seen that the last term is the angular momentum of the rigid body about its center of mass, in the frame of the center of mass. Therefore,

$$\boldsymbol{L}=\boldsymbol{L}_{CM}+M\boldsymbol{r}_{CM}\times\boldsymbol{v}_{CM}.$$

We see that the total angular momentum of the system is simply the sum of the angular momentum of the body about a point fixed to the center of mass and the angular momentum derived from treating the body as a point mass located at its center of mass and moving at $\boldsymbol{v}_{CM}$ with respect to the origin. Loosely put, the angular momentum has a rotational component and a translational component. To clarify this statement, translation commonly refers to the translation of the center of mass (which drags every other point with it while maintaining the orientation of the body) while rotation usually means the motion of the entire rigid body as viewed in the frame of the center of mass (so that the translational component is filtered out). Moving on, the z-component of the angular momentum is

$$L_z = I_{CM}\omega + M(\boldsymbol{r}_{CM}\times\boldsymbol{v}_{CM})_z.$$

The total kinetic energy can also be found to be

$$T=\frac{1}{2}\int(\boldsymbol{v}_{CM}+\boldsymbol{\omega}\times\boldsymbol{r}')\cdot(\boldsymbol{v}_{CM}+\boldsymbol{\omega}\times\boldsymbol{r}')dm$$

$$=\frac{1}{2}\int v_{CM}^2 dm+\frac{1}{2}\int(\boldsymbol{\omega}\times\boldsymbol{r}')\cdot(\boldsymbol{\omega}\times\boldsymbol{r}')dm+\boldsymbol{v}_{CM}\cdot\left(\boldsymbol{\omega}\times\int\boldsymbol{r}'dm\right).$$

The first term can be integrated trivially as v_{CM}^2 can be treated as a constant in the context of integrating over dm (v_{CM} may be changing over time but we are integrating over dm at a particular instance). From the expression for the kinetic energy in the previous section, it can be seen that the second term represents the total kinetic energy of the rigid body in the frame of the center of mass. Finally, the last term vanishes due to the definition of the center of mass. Therefore,

$$T = \frac{1}{2} I_{CM} \omega^2 + \frac{1}{2} M v_{CM}^2.$$

We can see that the total energy of the rigid body is simply the sum of the rotational kinetic energy of the body about its center of mass and the translational kinetic energy of the body at the velocity of the center of mass.

5.1.3 *Torque*

The torque $\boldsymbol{\tau}$, due to a force $\boldsymbol{F}$, with respect to a certain origin is defined as

$$\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F}, \tag{5.7}$$

where $\boldsymbol{r}$ is the position vector from the origin to the point where the force is exerted on. When acted upon a rigid body, a net torque leads to a rate of change of angular momentum — analogous to how a net force leads to a rate of change of linear momentum.[4] We shall now relate net torque — which is independent of the reference frame but dependent on the point we compute it about — to the angular momentum, which is dependent on the motion of the possibly accelerating origin.

The net external torque on a system about a certain origin is directly proportional to the rate of change of angular momentum of the system about the same origin — with the introduction of a fictitious $-M\ddot{\boldsymbol{r}}_0$ force[5] at the center of mass of the system if the origin is accelerating at $\ddot{\boldsymbol{r}}_0$.

$$\sum \boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}.$$

In our study of fixed axis rotations, the angular momentum vector does not change in direction. Therefore, $\frac{d\boldsymbol{L}}{dt} = \frac{dL}{dt}\hat{\boldsymbol{L}}$. Taking the z-component

[4]However, a net torque does not necessarily galvanize an angular acceleration, as opposed to the effect of a net force, which always leads to the acceleration of the center of mass. Consider the simple example where a block experiences a horizontal force exerted on the same vertical level as its center of mass on a frictionless ground. The torque about most origins is non-zero but the block merely translates without any rotation. The intuitive reason behind this lack of rotation is that the torque only contributes to the increase in the translational component of the angular momentum (associated with the motion of the center of mass) and not the rotational component.

[5]This is commonly known as the inertial force (see Chapter 11).

of both sides,

$$\sum \tau_z = \frac{dL_z}{dt}. \tag{5.8}$$

If we choose an origin at an ICoR (possibly identified after switching to the co-moving frame of a particle on the rigid body), $L_z = I\omega$ such that

$$\sum \tau = I\alpha, \tag{5.9}$$

where we have omitted the subscript z. For other choices of origins, we have to use the more general expression for angular momentum given by Eq. (5.5). Be wary that it is not always[6] the case that $\sum \tau = I\alpha$; $\boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}$ is always true but the exact relationship between torque and the angular acceleration depends on the choice of origin which determines the expression for the angular momentum. Lastly, in light of the fictitious force mentioned above, it is often ideal to fix the origin to the center of mass of an accelerating body to nullify the fictitious torque produced.

Proof: Notice that in the previous sections, the frames — in which angular momentum was calculated — are not necessarily inertial frames. However, Newton's laws will be a crucial part of the following analysis and we thus need to begin from an inertial frame with x, y and z-coordinate axes. Then, let the position vector of the origin O' — through which the angular momentum will be evaluated — in this inertial frame be $\boldsymbol{r}_0$ and the position vector of a point i on the body with respect to O' be $\boldsymbol{r}'_i = \boldsymbol{r}_i - \boldsymbol{r}_0$ where $\boldsymbol{r}_i$ is the position vector of point i in the inertial frame (Fig. 5.3).

Applying Newton's second law to the ith particle,

$$\left(\sum \boldsymbol{F}\right)_i + \sum_{j \neq i} \boldsymbol{f}_{ij} = m_i \ddot{\boldsymbol{r}}_i,$$

where $(\sum \boldsymbol{F})_i$ and $\boldsymbol{f}_{ij}$ refer to the net external force on the particle i and the internal force on particle j by another particle j respectively. The angular

[6] Actually, you can always apply $\sum \tau = I\alpha$ by fixing an origin to the rigid body. However, one then has to include the fictitious force associated with the acceleration of the origin which is often incapacitating as you generally do not know the acceleration of a particle on a rigid body (except for the center of mass), unless you have already solved the problem! Therefore, it is sometimes expeditious to choose a different origin that leads to a more cumbersome expression given by Eq. (5.5) to reap benefits in the long run. Other times, you can apply $\sum \tau = I\alpha$ about an origin attached to the center of mass of the **entire system**. If there are other discrete particles besides a rigid body, they must be considered too. Due to this reason, we usually compute the angular momenta of the rigid body and individual particles separately and then exploit the additive property of angular momentum to determine the total angular momentum.

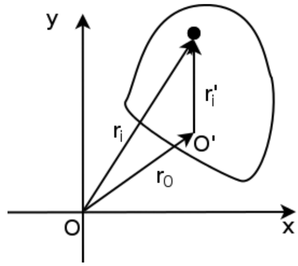


Figure 5.3: Possibly accelerating origin

momentum of the body with respect to O' is

$$\boldsymbol{L} = \sum_{i=1}^{n} \boldsymbol{r}'_i \times \boldsymbol{p}'_i = \sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times m_i(\dot{\boldsymbol{r}}_i - \dot{\boldsymbol{r}}_0).$$

Observe that the rate of change of angular momentum is

$$\begin{aligned}\frac{d\boldsymbol{L}}{dt} &= \sum_{i=1}^{n} (\dot{\boldsymbol{r}}_i - \dot{\boldsymbol{r}}_0) \times m_i(\dot{\boldsymbol{r}}_i - \dot{\boldsymbol{r}}_0) + \sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times m_i(\ddot{\boldsymbol{r}}_i - \ddot{\boldsymbol{r}}_0) \\ &= 0 + \sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times \left(\left(\sum \boldsymbol{F} \right)_i + \sum_{j \neq i} \boldsymbol{f}_{ij} - m_i \ddot{\boldsymbol{r}}_0 \right) \\ &= \sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times \left(\sum \boldsymbol{F} \right)_i + \sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times \left(\sum_{j \neq i} \boldsymbol{f}_{ij} \right) \\ &\quad - \left(\sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) m_i \right) \times \ddot{\boldsymbol{r}}_0.\end{aligned}$$

We can see that the first term is the net external torque with respect to O', $\sum \boldsymbol{\tau}$. The second term can be evaluated as follows by pairing the terms associated with $\boldsymbol{f}_{ij}$ with its conjugate $\boldsymbol{f}_{ji}$.

$$\begin{aligned}\sum_{i=1}^{n} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times \left(\sum_{j \neq i} \boldsymbol{f}_{ij} \right) &= \frac{1}{2} \sum_{i,j\, i \neq j} (\boldsymbol{r}_i - \boldsymbol{r}_0) \times \boldsymbol{f}_{ij} + (\boldsymbol{r}_j - \boldsymbol{r}_0) \times \boldsymbol{f}_{ji} \\ &= \frac{1}{2} \sum_{i,j\, i \neq j} (\boldsymbol{r}_i - \boldsymbol{r}_j) \times \boldsymbol{f}_{ij} \\ &= 0.\end{aligned}$$

The last equality assumes the strong law of action and reaction, which states that the action and reaction pair act along the line joining the two particles i

and j in addition to the fact that $\boldsymbol{f}_{ij} = -\boldsymbol{f}_{ij}$. In other words, $\boldsymbol{f}_{ij} \parallel (\boldsymbol{r}_i - \boldsymbol{r}_j)$. This assumption is valid in most cases — causing internal forces to produce no net torque.

In the limit where $n \to \infty$ and $m_i \to 0$, the third term becomes

$$\left(\int (\boldsymbol{r} - \boldsymbol{r}_0) dm\right) \times \ddot{\boldsymbol{r}}_0 = M(\boldsymbol{r}_{CM} - \boldsymbol{r}_0) \times \ddot{\boldsymbol{r}}_0,$$

where M is the total mass of the body and r_{CM} is the position vector of the center of mass as $\int \boldsymbol{r} dm = M\boldsymbol{r}_{CM}$ by definition. Combining these three terms, we obtain our general expression for $\frac{d\boldsymbol{L}}{dt}$.

$$\frac{d\boldsymbol{L}}{dt} = \sum \boldsymbol{\tau} - M(\boldsymbol{r}_{CM} - \boldsymbol{r}_0) \times \ddot{\boldsymbol{r}}_0.$$

In the case where O' accelerates at $\ddot{\boldsymbol{r}}_0$ and in light of the expression of the second term, we can actually imagine that the body experiences a fictitious force at the center of mass, given by $\boldsymbol{F} = -M\ddot{\boldsymbol{r}}_0$ and hence reduce the equation to $\frac{d\boldsymbol{L}}{dt} = \sum \boldsymbol{\tau}$ where $\sum \boldsymbol{\tau}$ includes the torque due to the fictitious force about the possibly-accelerating origin. Then,

$$\sum \boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}.$$

5.2 Moment of Inertia

The main objective of this section is to evaluate the moment of inertia with respect to a certain axis — a quantity given by

$$I = \int r_\perp^2 dm,$$

where $r_\perp$ is the perpendicular distance between an infinitesimal mass dm and the axis. The integration is performed over the entire mass distribution.

5.2.1 *Integration*

The most direct and general approach is to evaluate the integral. The crux is to, again, ensure that we actually integrate over the entire mass distribution.

Problem: Determine the moment of inertia of a uniform rod of mass m and length l about a perpendicular axis through its center.

Let the rod lie along the x-axis from $x = -\frac{l}{2}$ to $x = \frac{l}{2}$. The contribution of an infinitesimal element between x and $x + dx$ to the moment of inertia about an axis parallel to the x-axis and passing through the origin is $x^2dm = x^2\lambda dx$, where λ is the linear mass density of the rod. Hence,

$$I = \int_{-\frac{l}{2}}^{\frac{l}{2}} x^2\lambda dx$$
$$= \left[\frac{\lambda x^3}{3}\right]_{-\frac{l}{2}}^{\frac{l}{2}}$$
$$= \frac{1}{12}\lambda l^3.$$

Since $m = \lambda l$,

$$I = \frac{1}{12}ml^2.$$

Problem: Determine the moment of inertia of a uniform solid circle of mass M and radius R about a perpendicular axis through its center.

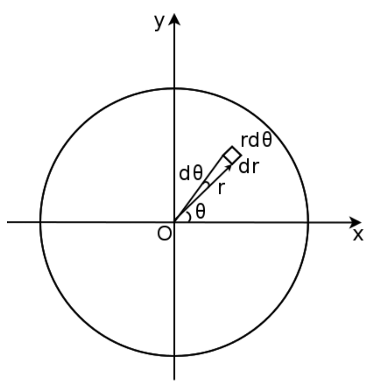


Figure 5.4: Circle

Consider an infinitesimal element $dm = \sigma r d\theta dr$ in polar coordinates (Fig 5.4), where σ is the surface mass density. Its contribution to the moment of inertia about the z-axis is $r^2dm = r^3\sigma d\theta dr$. Next, we have to integrate this quantity over the entire circle — with θ ranging from 0 to 2π and r from 0 to R.

$$I = \int_0^R \int_0^{2\pi} r^3\sigma d\theta dr$$
$$= \sigma \int_0^R 2\pi r^3 d\theta dr$$
$$= \frac{1}{2}\sigma\pi R^4.$$

Since $\sigma\pi R^2 = M$,

$$I = \frac{1}{2}MR^2.$$

Problem: Determine the moment of inertia of a uniform solid sphere of mass M and radius R about an axis through its center.

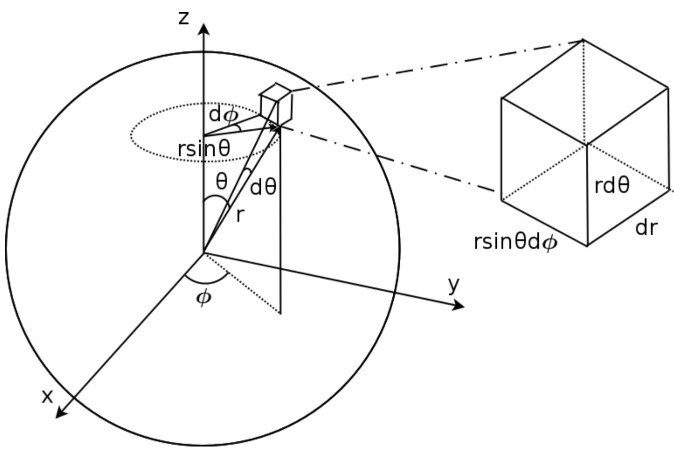


Figure 5.5: Spherical coordinates

In spherical coordinates, the infinitesimal mass element at (r, ϕ, θ) is a box with $dm = \rho r^2 \sin\theta d\phi dr d\theta$, where ρ is the mass density. The perpendicular distance between an infinitesimal element and the z-axis is $r_\perp = r\sin\theta$.

$$\begin{aligned} I &= \int r_\perp^2 dm = \int_0^\pi \int_0^R \int_0^{2\pi} r^2 \sin^2\theta \rho r^2 \sin\theta d\phi dr d\theta \\ &= 2\rho\pi \int_0^\pi \int_0^R r^4 \sin^3\theta dr d\theta = \frac{2}{5}\rho\pi R^5 \int_0^\pi \sin^3\theta d\theta \\ &= \frac{8}{15}\rho\pi R^5 = \frac{2}{5}MR^2. \end{aligned}$$

The integral $\int \sin^3\theta d\theta$ can be integrated by the procedure: $\int(1 - \cos^2\theta)\sin\theta d\theta = -\cos\theta + \int u^2 du = -\cos\theta + \frac{\cos^3\theta}{3} + c$ where substitutions $u = \cos\theta$ and $du = -\sin\theta d\theta$ have been made.

5.2.2 *Parallel Axis Theorem*

Suppose that we know the moment of inertia of an object about an axis through its center of mass, I_{CM}, and wish to determine its moment of inertia

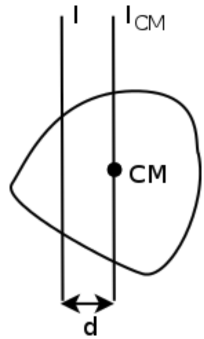


Figure 5.6: Parallel axis theorem

about another parallel axis, I. The parallel axis theorem states that

$$I = I_{CM} + md^2,$$

where d is the distance between the two axes and m is the total mass of the mass distribution (Fig. 5.6). This is valid for all parallel axes, including those outside of the object.

Proof: Let the origin of the coordinate system be at the center of mass of the distribution and define the z-axis to be the axis through which the moment of inertia is computed. Then, let the second axis be located $x = x_0$, $y = y_0$. With these definitions,

$$\begin{aligned}
I_{CM} &= \int (x^2 + y^2) dm \\
I &= \int (x - x_0)^2 + (y - y_0)^2 dm \\
&= \int (x^2 + y^2) dm - 2x_0 \int x dm - 2y_0 \int y dm + \int (x_0^2 + y_0^2) dm \\
&= I_{CM} + 0 + 0 + m(x_0^2 + y_0^2) \\
&= I_{CM} + md^2.
\end{aligned}$$

The two zeroes arise from the definition of the center of mass which is located at the origin — implying that $\int x dm = 0$ and $\int y dm = 0$.

Problem: Determine the moment of inertia of a uniform solid circle of mass M and radius R about a perpendicular axis through its circumference.

Applying the parallel axis theorem with $d = R$,

$$I_{circum} = I_{CM} + MR^2 = \frac{3}{2}MR^2.$$

5.2.3 *Perpendicular Axis Theorem*

For a pancake (flat) object, the moment of inertia through an axis normal to the plane of the object — defined to be the z-axis — is equal to the sum of the moments of inertia about two mutually-perpendicular axes in the plane of the object. All of these axes are concurrent at a point O.

$$I_x + I_y = I_z.$$

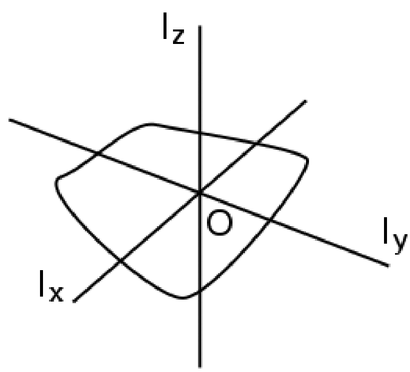


Figure 5.7: Perpendicular axis theorem

Proof: Let the axis normal to the plane be z and those in the plane be x and y. Since the object is flat, the z-coordinates of its elements are zero if the origin is defined at the point of concurrency.

$$I_x = \int (y^2 + z^2)dm = \int y^2 dm,$$

$$I_y = \int (x^2 + z^2)dm = \int x^2 dm,$$

$$I_x + I_y = \int (x^2 + y^2)dm = I_z.$$

Problem: Determine the moment of inertia of a uniform solid circle of mass M and radius R about a diameter.

Let the required moment of inertia be evaluated about the x-axis. Due to the rotational symmetry of the circle, $I_y = I_x$ where the y-axis lies in the plane of the circle and is perpendicular to the x-axis. Then,

$$I_x + I_y = I_z$$
$$2I_x = \frac{1}{2}MR^2$$
$$I_x = \frac{1}{4}MR^2.$$

5.2.4 *Squashing*

In light of the limited applicability of the perpendicular axis theorem to merely pancake objects, it would be ideal if we could reduce an n-dimensional object ($n = 2$ or 3) to an equivalent $(n-1)$-dimensional object with the same moment of inertia.

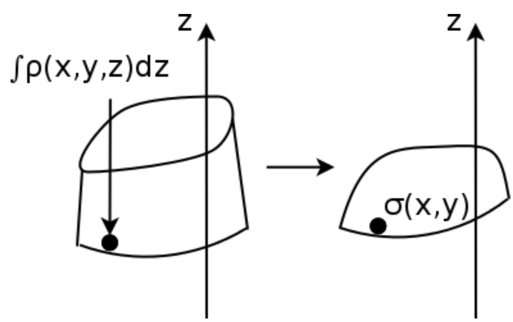


Figure 5.8: Squashing in the z-direction

Suppose that we wish to determine the moment of inertia of a three-dimensional object about the z-axis.

$$I = \int (x^2 + y^2)dm = \int (x^2 + y^2)\rho(x, y, z)dxdydz,$$

where $\rho(x, y, z)$ is the mass density of the distribution. Notice that we can first perform the integration over z such that

$$\begin{aligned} I &= \iiint (x^2 + y^2)\rho(x, y, z)dzdxy \\ &= \iint (x^2 + y^2) \cdot \left(\int \rho(x, y, z)dz \right) dxdy \\ &= \iint (x^2 + y^2) \cdot \sigma(x, y)dxdy, \end{aligned}$$

with $\sigma(x,y) = \int \rho(x,y,z)dz$ being the surface mass density of the equivalent two-dimensional object. Referring to Fig. 5.8, this operation effectively squashes a three-dimensional object such that all masses along an axis at a certain (x,y), parallel to the z-axis, collapse to a single point for all (x,y). Note that the total mass is still preserved.

A direct corollary of this is that the moment of inertia of a three-dimensional object with a uniform cross section[7] about the z-axis is equivalent to that of its cross section with an identical total mass.

Problem: Determine the moment of inertia of a uniform solid cylinder of mass m, radius r and length l about its cylindrical axis.

We can squash the cylinder along its cylindrical axis into a uniform circle. Hence, the required moment of inertia is $\frac{1}{2}mr^2$. It is easy to visualize the effect of "squashing" via direct integration. In cylindrical coordinates, the required integral is (with ρ being the mass density and r being the radial distance perpendicular to the cylindrical axis)

$$I = \int_0^R \int_0^{2\pi} \int_0^l \rho r^3 dz d\theta dr$$
$$= \int_0^R \int_0^{2\pi} \rho l r^3 d\theta dr.$$

At this point, we can stop and observe that this expression is equivalent to that of a circle with surface mass density $\sigma = \rho l$.

For certain mass distributions, a combination of "squashing" and the perpendicular axis theorem can prove to be potent in finding their moment of inertia.

Problem: Determine the moment of inertia of a uniform solid cylinder of mass m, radius r and length l about an axis that is perpendicular to its cylindrical axis, passing through its center of mass.

We can first squash the cylinder along the z-axis to obtain a non-uniform mass distribution in the shape of a rectangle in the xy-plane (Fig. 5.9).

The moment of inertia of this non-uniform plate is tedious to determine. However, we can now apply the perpendicular axis theorem

$$I_z = I_x + I_y.$$

[7] Uniform in the sense that cross sections at different z-coordinates are identical.

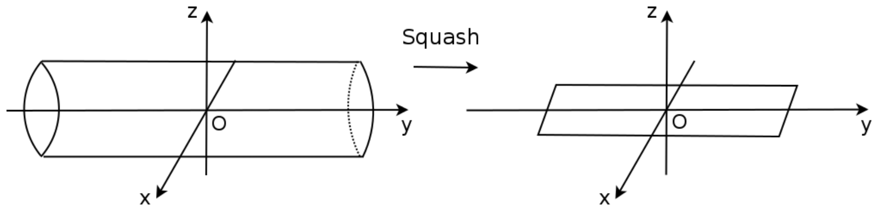


Figure 5.9: Squashing a cylinder into a plane

To compute I_x and I_y, the plate can be squashed along the corresponding directions again to obtain a one-dimensional distribution. In the x-direction, the squashed plate becomes a uniform rod of length l and mass m. Hence,

$$I_x = \frac{1}{12}ml^2.$$

The moment of inertia of the resultant distribution (the non-uniform rod in the middle of Fig. 5.10) after squashing in the y-direction is less obvious. However, we can leverage symmetry arguments to assert that the moments of inertia I'_y and I'_z of the resultant rod, relative to the y and z-axes, are identical.

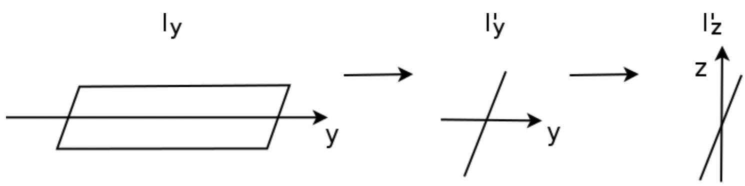


Figure 5.10: Squashing the plane in the y-direction

$$I'_y = I'_z.$$

Since the rod is equivalent to a uniform circle — of radius r and mass m in the x-z plane — squashed in the z-direction, the required moment of inertia is equal to that of the uniform circle about a diameter.

$$I'_z = \frac{1}{4}mr^2.$$

Therefore,

$$I_z = I_x + I_y = \frac{1}{12}ml^2 + \frac{1}{4}mr^2.$$

5.2.5 *Moment of Inertia About a Slanted Axis*

In some cases, the moment of inertia about a non-conventional axis A may be required. Then, the $r_\perp^2$ with respect to A should be expressed in terms of the coordinates associated with the conventional axes x, y and z so that this moment of inertia can be related to those about the conventional axes. All axes pass through the same origin. The most general method of determining $r_\perp^2$ involves the use of vectors.[8]

Let $\hat{\boldsymbol{d}}$ be the direction vector of axis A, which can be expressed in x, y and z coordinates. Let the position vector of an infinitesimal mass element under consideration be $\boldsymbol{r}$. Then, the magnitude of the component of $\boldsymbol{r}$ parallel to $\hat{\boldsymbol{d}}$ is $\boldsymbol{r} \cdot \hat{\boldsymbol{d}}$. The squared magnitude of the component perpendicular to axis A is then obtained from Pythagoras' theorem.

$$r_\perp^2 = |\boldsymbol{r}|^2 - (\boldsymbol{r} \cdot \hat{\boldsymbol{d}})^2.$$

Problem: Determine the moment of inertia of a rectangle, with side lengths a and b, about axis A in the plane of the rectangle as shown in Fig. 5.11.

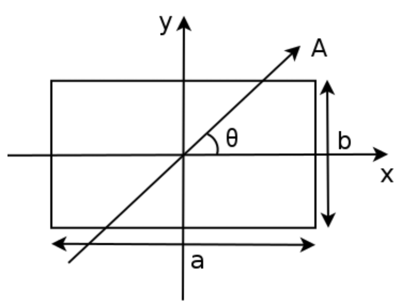


Figure 5.11: Moment of inertia of rectangle about slanted axis

$$\hat{\boldsymbol{d}} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix},$$

$$\boldsymbol{r} = \begin{pmatrix} x \\ y \end{pmatrix},$$

$$\boldsymbol{r} \cdot \hat{\boldsymbol{d}} = x\cos\theta + y\sin\theta,$$

$$r_\perp^2 = x^2 + y^2 - (x\cos\theta + y\sin\theta)^2 = x^2\sin^2\theta + y^2\cos^2\theta - 2xy\sin\theta\cos\theta,$$

[8] Alternatively, one can also use the rotation matrix.

$$I = \int r_\perp^2 dm$$
$$= \sin^2\theta \int x^2 dm + \cos^2\theta \int y^2 dm - 2\sin\theta\cos\theta \int xydm$$
$$= \sin^2\theta I_y + \cos^2\theta I_x - 0.$$

The integral $\int xydm$ evaluates to zero as the distribution is symmetric — for every point (x, y), there is a counterpart at $(-x, y)$. Now, I_x and I_y can be determined by squashing the rectangle into rods along the corresponding axes to obtain

$$I_x = \frac{1}{12}mb^2,$$
$$I_y = \frac{1}{12}ma^2.$$

Therefore,

$$I = \frac{1}{12}ma^2\sin^2\theta + \frac{1}{12}mb^2\cos^2\theta.$$

5.2.6 *Scaling Arguments*

Observe that the moment of inertia of an object about an axis is strictly proportional to the mass of the object and the square of its length dimensions. If we scale all length dimensions of an n-dimensional object ($n = 1, 2$ or 3) by a factor of k, the moment of inertia will vary by a factor k^2 due to the change in length dimensions and another factor associated with the change in the mass of the scaled object. The latter factor depends on the configuration of the actual mass distribution and, usually, its number of dimensions (for a distribution with no holes, the latter factor is k^n such that the overall factor is k^{2+n}).

If the object can be conveniently related to smaller versions of itself, we can use a scaling argument to calculate its moment of inertia via the parallel axis theorem — precluding any need for integration.

Problem: Calculate the moment of inertia of a rod of mass m and length l along an axis passing through its center of mass.

Referring to Fig. 5.12, the first equation argues that the moment of inertia of a rod of length $2l$ will be 8 times that of a rod of length l, both measured through their center of masses. This is true because $I \propto ml^2$ and both the mass and length of the rod doubles. The second equation argues that the moment of inertia of a rod of length $2l$ through its center of mass is two

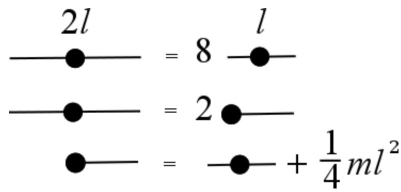


Figure 5.12: Scaling argument for rod

times the moment of inertia of a rod of length l through one of its ends. This can be obtained if you imagine that you break the longer rod into two at the center of mass. Lastly, the third equation is obtained by applying the parallel axis theorem. Solving the simultaneous equations above, we obtain the moment of inertia for a rod of length l about a origin at the center of mass,

$$I_{CM} = \frac{1}{12}ml^2.$$

Problem: Take an equilateral triangle of length l and remove an equilateral triangle of half its edge length from its center. Then, repeat this process for the three equilateral triangles of halved length dimensions formed and so on. If the mass of the resultant fractal is m, determine its moment of inertia about a perpendicular axis through its center of mass.

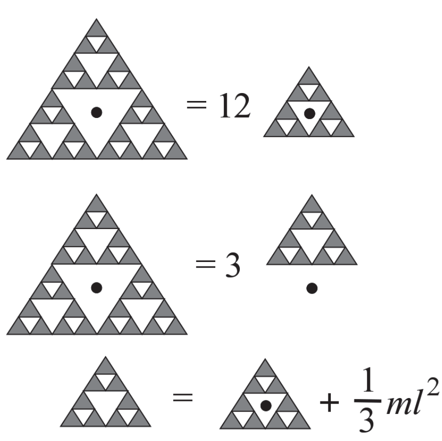


Figure 5.13: Scaling argument for fractal

Referring to Fig 5.13 the first equation asserts that the moment of inertia of a fractal with length $2l$ is 12 times that of the original fractal as it has triple its mass — evident from the fact that it is composed of three original fractals — and double its length dimensions ($3 \times 2^2 = 12$). The second equation simply decomposes the moment of inertia of a fractal with length $2l$ into the moment of inertia of three fractals about an off-center axis. The third equation is an application of the parallel axis theorem (the distance between the two axes can be shown to be $\frac{\sqrt{3}}{3}l$). Solving the above simultaneous equations, the relevant moment of inertia is

$$I = \frac{1}{9}ml^2.$$

For more examples of scaling arguments, refer to Ref. [1].

5.3 Applying $\boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}$

We shall practise applying $\boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}$ to a few systems — in this process, keep in mind that $\sum \boldsymbol{F}_{ext} = M\boldsymbol{a}_{CM}$ for a macroscopic body (see Chapter 4). Note that since the states of all of the following problems can be quantified by a single coordinate each, their velocities (possibly angular) can be solved from the conservation of energy directly while their accelerations can be determined by differentiating the conservation of energy equation. The reader should try to analyze these systems via the conservation of energy as an exercise.

Problem: A mass m is hung over a wheel of radius R. The wheel cannot translate but it can rotate freely. Let the moment of inertia of the wheel be I with respect to an longitudinal axis through its center. Find the acceleration of the mass, assuming that the string remains taut at all times.

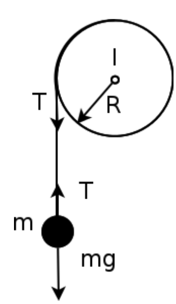


Figure 5.14: Wheel and mass

Such problems can often be solved in two different ways by considering two different systems. One way accounts for the internal forces between different objects and applies $\tau = I\alpha$ and $F = ma$ to all objects. The other method uses $\tau = \frac{dL}{dt}$ for a larger system comprising various objects. The former is often more straightforward while the latter is terser as it does not consider internal forces, but the direction of the components of angular momentum can be rather confusing. We shall illustrate both here.

Method 1: We use the axle as our origin as we do not know the normal force exerted by the axle on the wheel and have no wish to. Let the tension in the string be T. Considering the torque on the wheel about the origin while taking the anti-clockwise direction to be positive,

$$\tau = TR.$$

The rate of change of the angular momentum of the wheel about its center is $I\alpha$ as the center is an ICoR. Thus,

$$I\alpha = TR.$$

Considering the forces on the mass,

$$ma = mg - T$$

where the positive direction has been assumed to be downwards. Furthermore, by the conservation of string,

$$a = R\alpha.$$

Solving the equations above,

$$a = \frac{mgR^2}{I + mR^2}.$$

Method 2: Considering the torques and angular momentum of the whole system about the axle,

$$\frac{dL}{dt} = \frac{d(L_{wheel} + L_{mass})}{dt} = \frac{d(I\omega + mvR)}{dt} = I\alpha + maR,$$

where $L_{wheel} = I\omega$ (as the center is an ICoR) and $L_{mass} = mvR$ (angular momentum of a discrete particle). The only contributor to the torque about the center is the weight of the mass acting at a perpendicular distance R

away from the origin.

$$\tau = mgR$$

$$mgR = I\alpha + maR = \left(\frac{I}{R} + mR\right)a,$$

as $a = R\alpha$ by the conservation of string. Simplifying,

$$a = \frac{mgR^2}{I + mR^2}.$$

Problem: Referring to Fig. 5.15, a uniform sphere, of mass m, radius R and moment of inertia I about an axis through its center, begins accelerating down an inclined plane with an angle of inclination, θ. Find the acceleration of the center of the sphere, assuming that it rolls without slipping and that the plane remains stationary.

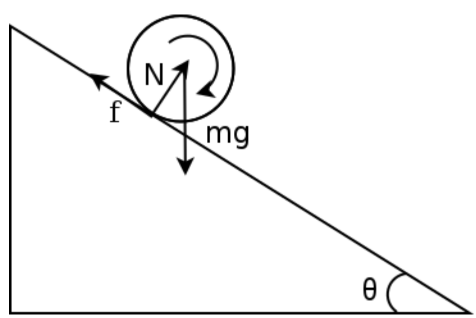


Figure 5.15: Sphere on inclined plane

Method 1: Define a_{CM} to be the acceleration of the center of the sphere, parallel to the slope and positive downwards. Writing $F = ma$ for the sphere and $\tau = I\alpha$ about an origin fixed to its center,

$$mgsin\theta - f = ma_{CM},$$

$$fR = I\alpha, \qquad \text{(About the center)}$$

$$a_{CM} = R\alpha. \qquad \text{(Non-slip condition)}$$

Solving,

$$a_{CM} = \frac{mgR^2 sin\theta}{I + mR^2}.$$

Method 2: Define the origin at a point on the surface of the plane such that it is stationary. The clockwise angular momentum of the sphere is then given by Eq. (5.5) as $I\omega + mv_{CM}R$ where v_{CM} is the velocity of the center of mass down the slope. Accordingly,

$$\tau = \frac{dL}{dt} \implies mgsin\theta R = \frac{d(I\omega + mv_{CM}R)}{dt} = I\alpha + ma_{CM}R,$$

$$a_{CM} = R\alpha, \qquad \text{(Non-slip condition)}$$

$$a_{CM} = \frac{mgR^2 sin\theta}{I + mR^2}.$$

Note that we could have chosen a third origin — the instantaneous point of contact between the sphere and the slope. Then, $L = I_{surface}\omega = (I + mR^2)\omega$ where $I_{surface}$ is the moment of inertia of the sphere about the point of contact which functions as an ICoR (it is stationary by the non-slip condition). It does not matter if we fix the origin to the sphere or choose it on the slope as the point of contact on the sphere does not possess an instantaneous tangential acceleration anyway (it has a centripetal acceleration but the fictitious force associated with this produces no torque about the point of contact). Though we have to continuously adjust the origin due to the perpetually changing nature of the point of contact, $\tau = \frac{dL}{dt} = \frac{d(I_{surface}\omega)}{dt}$ is always instantaneously true and can thus be applied to find the instantaneous accelerations which are germane.

Problem: A uniform ladder, of mass m and length $2l$, is initially held motionless, parallel to the wall. It is then given a slight push and released. The top end begins to slide down the wall whereas the bottom end slides along the ground. Assuming that all surfaces are frictionless, find $\ddot{\theta}(\theta)$, $\dot{\theta}(\theta)$ and the angle θ at which the ladder loses contact with the wall.

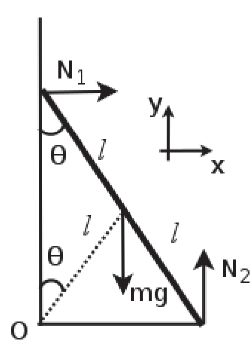


Figure 5.16: Ladder on wall

To solve the first part of the problem, we write our $F = ma$ and $\tau = \frac{dL}{dt}$ equations. Since points on the ladder are generally accelerating, we choose the center of mass as the point with respect to which we calculate angular momentum and torques, for the sake of convenience.

$$m\ddot{x}_{CM} = N_1,$$

$$m\ddot{y}_{CM} = N_2 - mg,$$

$$\tau = N_2 l \sin\theta - N_1 l \cos\theta = \frac{dL}{dt} = I_{CM}\ddot{\theta} = \frac{1}{3}ml^2\ddot{\theta}.$$

Note that θ increases in the anti-clockwise direction, therefore it is taken to be positive when writing the above equation. Now we have 5 variables — $\ddot{x}_{CM}$, $\ddot{y}_{CM}$, $\ddot{\theta}$, N_1 and N_2 — but only three equations. We can obtain the remaining two by relating $\ddot{x}_{CM}$, $\ddot{y}_{CM}$ and $\ddot{\theta}$ via the observations that the ends of the ladder must stick to the wall and ground and that the length of the ladder is fixed.

Define the origin of our fixed coordinate system at the bottom of the wall. Then, the coordinates of the center of mass of the ladder can be related to angle θ.

$$x_{CM} = l\sin\theta,$$

$$y_{CM} = l\cos\theta,$$

$$\dot{x}_{CM} = l\cos\theta\dot{\theta},$$

$$\dot{y}_{CM} = -l\sin\theta\dot{\theta},$$

$$\ddot{x}_{CM} = -l\sin\theta\dot{\theta}^2 + l\cos\theta\ddot{\theta},$$

$$\ddot{y}_{CM} = -l\cos\theta\dot{\theta}^2 - l\sin\theta\ddot{\theta}.$$

Solving the five equations above yields

$$\ddot{\theta} = \frac{3g\sin\theta}{4l}.$$

$\dot{\theta}$ is obtained from separating variables in $\ddot{\theta} = \dot{\theta}\frac{d\dot{\theta}}{d\theta}$ and integrating.

$$\int_0^{\dot{\theta}} \dot{\theta} d\dot{\theta} = \int_0^{\theta} \frac{3g\sin\theta}{4l} d\theta$$

$$\dot{\theta} = \sqrt{\frac{3g}{2l}(1-\cos\theta)}.$$

Actually, this result, which originated from such a complicated process, can readily be obtained from the conservation of energy, as we shall see in the next chapter. Back to the main point, since $N_1 = m\ddot{x}_{CM}$, the ladder loses contact with the wall when $\ddot{x}_{CM} = 0$. From the expression for $\ddot{x}_{CM}$ above, this occurs when

$$l \sin\theta \cdot \frac{3g}{2l}(1 - \cos\theta) = l\cos\theta \cdot \frac{3g\sin\theta}{4l}$$

$$\theta = \cos^{-1}\frac{2}{3},$$

where the uneventful $\sin\theta = 0$ solution has been rejected.

Problem: A uniform stick of length $2l$ and mass m is held motionless along the vertical initially and is then gently released on a frictionless, horizontal ground. Obtain an equation that relates the angular acceleration $\ddot{\theta}$, angular velocity $\dot{\theta}$ and θ, the angle between the stick and the vertical. From there, use the method of integrating factors to obtain $\dot{\theta}$ in terms of θ.

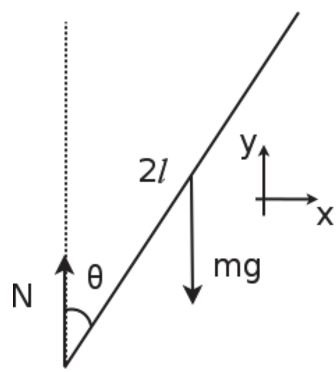


Figure 5.17: Toppling stick

In the previous accelerating set-ups, the origin was always chosen at the center of mass. Through this problem, we will underscore the importance of choosing a convenient origin to calculate our angular momentum and $\frac{dL}{dt}$ by venturing into accelerating origins that are not the center of mass. We will solve this part twice — using two different origins. We first write our $F = ma$ equations.

$$m\ddot{x}_{CM} = 0,$$

$$N - mg = m\ddot{y}_{CM} = -m(l\cos\theta\ddot{\theta}^2 - l\sin\theta\ddot{\theta}).$$

Origin 1: We fix our origin at the center of mass of the stick. Then,

$$\tau = Nl\sin\theta.$$

Because we have defined our origin at the center of mass, even though the stick is generally accelerating,

$$\frac{dL}{dt} = I_{CM}\ddot{\theta} = \tau$$

$$\frac{1}{3}ml^2\ddot{\theta} = Nl\sin\theta$$

$$N = \frac{ml}{3\sin\theta}\ddot{\theta}.$$

Substituting this back into the second equation,

$$\cos\theta\dot{\theta}^2 + \frac{3\sin^2\theta + 1}{3\sin\theta}\ddot{\theta} = \frac{g}{l}.$$

To solve this equation, use the trick $\ddot{\theta} = \frac{1}{2}\frac{d\dot{\theta}^2}{d\theta}$. Then,

$$\cos\theta\dot{\theta}^2 + \frac{3\sin^2\theta + 1}{6\sin\theta}\frac{d\dot{\theta}^2}{d\theta} = \frac{g}{l}.$$

Multiplying by the integrating factor $6\sin\theta$,

$$6\sin\theta\cos\theta\dot{\theta}^2 + (3\sin^2\theta + 1)\frac{d\dot{\theta}^2}{d\theta} = \frac{d\left[(3\sin^2\theta + 1)\dot{\theta}^2\right]}{d\theta} = \frac{6g}{l}\sin\theta.$$

Separating variables and integrating,

$$\int_0^{(3\sin^2\theta+1)\dot{\theta}^2} d\left[(3\sin^2\theta + 1)\dot{\theta}^2\right] = \int_0^{\theta}\frac{6g}{l}\sin\theta d\theta$$

$$\dot{\theta} = \sqrt{\frac{6g(1-\cos\theta)}{l(3\sin^2\theta + 1)}}.$$

This equation could have, again, been obtained directly from the conservation of energy — bearing credence to its utility.

Origin 2: The bottom end of the stick is seemingly a good origin to pick as the unknown normal force acts at that point. However, we will discover that the ensuing process is in fact much more tedious. Fixing the origin at the bottom end and calculating our external torques,

$$\tau = mgl\sin\theta.$$

As our origin is now accelerating, we must introduce another fictitious component of torque due to a fictitious horizontal force $-m\ddot{x}_{end}$ at the center of mass where x_{end} is the horizontal coordinate of the end of the ground. If we define the $x = 0$ point at the center of mass which does not translate in the x-direction (as there are no horizontal forces),

$$x_{end} = -l \sin\theta$$
$$-\ddot{x}_{end} = -l \sin\theta \dot{\theta}^2 + l \cos\theta \ddot{\theta}.$$

The fictitious force $-m\ddot{x}_{end}$ thus introduces a fictitious torque, about the bottom end of the stick, given by

$$\tau_{fict} = ml \cos\theta (l \cos\theta \ddot{\theta} - l \sin\theta \dot{\theta}^2).$$

The relationship between the torques and the rate of change of angular momentum, all about the end of the stick is

$$\frac{dL}{dt} = \tau + \tau_{fict}.$$

As our origin is fixed to the rotating body,

$$L = I_{end}\omega = (I_{CM} + ml^2)\dot{\theta}$$
$$\frac{4}{3}ml^2\ddot{\theta} = mgl \sin\theta + ml \cos\theta (l \cos\theta \ddot{\theta} - l \sin\theta \dot{\theta}^2).$$

After some simplification,

$$\cos\theta \dot{\theta}^2 + \frac{3\sin^2\theta + 1}{3\sin\theta}\ddot{\theta} = \frac{g}{l},$$

which is the same equation as before. We can then proceed to solve the differential equation. These two different methods emphasize the importance of choosing a convenient origin. For the second origin, we may have skimped on the calculation of the real torques but we spent significantly more effort in relating $\frac{dL}{dt}$ to τ with the addition of a fictitious torque τ_{fict}.

As illustrated by this problem, the general rule of thumb in rotational dynamics is to choose the center of mass as the origin — regardless of how tempting other origins might be — as it eradicates the need to consider fictitious torques. This is contrary to the stance in statics which encourages the judicious choice of origins to eliminate as many irrelevant forces as possible. However, the latter is precisely possible because we are guaranteed the absence of accelerating origins which precludes fictitious torques. Furthermore, we are not obliged to relate the torque to any property of a static system (i.e. we only need to solve $\sum \boldsymbol{\tau} = \mathbf{0}$ and not $\sum \boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}$) so our fears of choosing an origin that complicates the calculation of $\boldsymbol{L}$ can also be dispelled.

Problems

Moment of Inertia

1. *Circular Hoop**

Determine the moment of inertia of a thin circular hoop of radius r and mass m about a perpendicular axis through its center of mass.

2. *Spherical Shell**

Determine the moment of inertia of a spherical shell of total mass m and radius r about an axis through its center.

3. *Cube***

Find the moment of inertia of a cube of length l, mass m and uniform mass density about an axis through its center of mass and perpendicular to two of its faces using: 1) Integration 2) Squashing and Perpendicular Axis Theorem 3) Scaling Arguments.

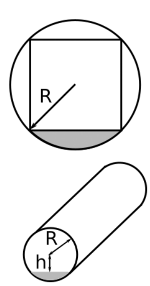
4. *Cone***

Determine the moment of inertia of a cone of half-angle θ, mass m and height l about its height.

5. *Cylinder with Liquid***

Find the moment of inertia of the shaded region, with uniform mass density σ, about an axis passing perpendicularly through the center of the circle in the top figure.

Suppose that you fill a hollow cylinder of radius R and length L with some liquid of mass density ρ as shown in the bottom figure (only the side view of the liquid is shown). Determine the moment of inertia of the liquid about the axis of the cylinder. Verify your answers for the limiting cases: $h = 0$, $h = R$, $h = -R$ and $h = \frac{R}{\sqrt{2}}$.



6. *Cube about Any Axis*******

Prove that the moment of inertia of a uniform cube of mass m and length l is

$$I = \frac{1}{6}ml^2$$

about any axis passing through its center of mass.

7. *Cube Fractal*******

Suppose you have a cube of length l. You cut the cube into 27 $\frac{l}{3} \times \frac{l}{3} \times \frac{l}{3}$ cubes and remove the middle cube. You then repeat this process for each of the subsequent 26 cubes and so on. If the total mass of the resultant fractal is m, determine the moment of inertia of the resultant fractal about an axis through its center of mass which is perpendicular to a face of the original cube.

8. *Cone about Slanted Axis********

Determine the moment of inertia of a cone with half-angle θ and height l about an axis A passing through its vertex and tangential to its surface (i.e. the axis and the height of the cone subtends an angle θ).

Torque and Angular Momentum

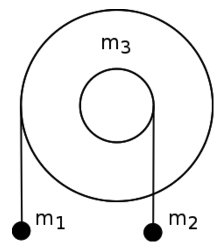
9. *Angular Momentum of a Particle*****

The expression for the angular momentum of a point mass about an origin is given by $\boldsymbol{L} = 2\alpha t \ln(\frac{t}{t+1})\hat{\boldsymbol{k}}$ where α is a certain constant.

a) Find the expression for the torque acting on the mass as a function of time.
b) Knowing that $\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F}$, show that the mass will never reach point $(0, 0, 2)$.
c) If the position of the particle at $t = 1$ is $(0, 2, 0)$, find the x and z components of its momentum at $t = 1$.

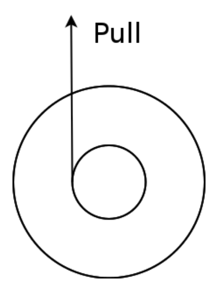
10. *Massive Atwood's Machine*****

A wheel of outer radius r_1 is wrapped around an axle of radius r_2. The wheel's thickness is equal to the length of the axle and the wheel and axle have the same uniform mass density. If the combined mass of the wheel and axle is m_3 and two masses m_1 and m_2 hang vertically from the wheel and axle respectively, determine the angular acceleration of the wheel. The axle cannot translate and does not slip with respect to the wheel.



11. ***Yo-yo****

A massless string is wound around a circular axle of radius r. A wheel of radius R is wrapped around the axle. If the moment of inertia of the wheel-cum-axle about its center is $\frac{1}{2}mR^2$ where m is their total mass and the wheel and axle do not slip, determine the angular acceleration of the wheel-cum-axle. Assume that you hold the string vertically such that the string remains taut.



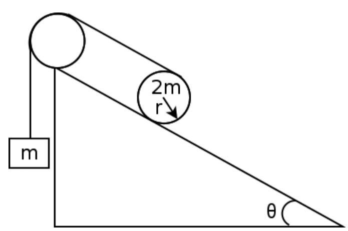
12. ***Rotating Ball*****

In all of the following set-ups, the center of a uniform sphere of radius r and mass m undergoes uniform circular motion about an origin O on a horizontal table at radius R and anti-clockwise angular velocity Ω. In the first set-up, the ball slides at a constant (center of mass) speed while not spinning at all. In the second set-up, the ball slides at a constant speed but also spins at a constant angular velocity about an axis perpendicular to the table such that the right-most (with an arbitrary definition of right in the plane of the table) side of the ball is the same side throughout. Finally, in the third set-up, the ball is rolling without slipping with a constant (center of mass) speed and angular speed on the table. Note that when we refer to spinning at a certain angular velocity, we mean the angular velocity that is observed when we switch to the rotating frame with coordinate axes fixed to the center of mass of the ball.

Determine the angular momentum, with respect to a point that is a height r above the origin O, and the kinetic energy of the ball in the lab frame in all three situations. In which case(s) does the ball experience a net external torque about its center of mass? Try to determine the net external torque(s).

13. *Another Atwood***

A cylinder of mass $2m$ and radius r is hung over a pulley and connected to another weight of mass m on a massive inclined plane with an angle of inclination θ. The string is wrapped around the cylinder and the free segment is tangential to its top. Assuming that there is no slipping between all surfaces and that the string remains taut throughout the motion, find the angular acceleration of the cylinder.



14. *Staying on a Ramp***

A cylinder of mass m and radius r rolls without slipping on an initially horizontal ground. It then encounters a slope of angle α below the horizontal. What is the maximum center of mass velocity v_0 for which the cylinder will not lose contact with the surface? Try using the conservation of energy, which will be introduced in Chapter 6.

15. *Spinning Disc***

A uniform circular disc of radius R is spinning about its vertical axis on a rough table. If its initial angular speed is ω_0 and the coefficient of kinetic friction between the table and the disc is μ, determine the time it takes for the disc to come to rest.

16. *Spinning Fan***

A fan hangs vertically from a ceiling. The fan is composed of four identical, uniform and rectangular blades of mass m, width a (the width is aligned with the vertical) and length b connected to a central axis. If the drag force per unit area is Dv^2 where v is the speed of the area of concern, determine the angular velocity of the fan $\omega(t)$ if its initial value is ω_0.

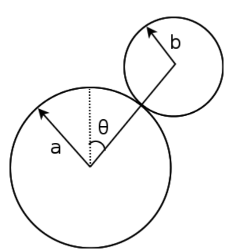
17. *Sweeping Rod***

A rod of mass m and length l is pivoted at one of its ends on a horizontal table. If the table is covered with dust with a uniform surface mass density σ, determine the minimum amount of force required to keep the rod moving at a constant angular velocity ω as it rotates one round about the pivot.

18. *Sphere and Cylinder***

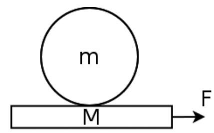
A sphere of radius b, mass m and uniform mass density initially lies motionless on top of a cylinder of radius a. It then begins to roll without slipping on the exterior of the cylinder.

a) What is the condition for the sphere to not slip from the cylinder? Express this in terms of $\ddot{\theta}$ and the angular acceleration of the sphere about its center.
b) Find $\ddot{\theta}$ as a function of θ.
c) Find $\dot{\theta}$ as a function of θ.
d) For arbitrarily large coefficients of static friction, find the angle θ_0 at which the sphere is guaranteed to lose contact with the cylinder.



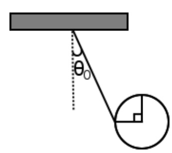
19. *Sphere on Block***

A sphere of radius r, mass m and uniform mass density lies on top of a block of mass M. The block lies on a horizontal, frictionless table. If the coefficient of static friction between the block and the sphere is μ, determine the largest force F exerted on the block such that the sphere does not slip relative to the block. In this case, under what conditions does the center of the sphere move forwards and backwards respectively?



20. *Pendulum**

A uniform sphere of mass m and radius r is hung from a massless string of length l that currently makes an angle $\theta = \theta_0$ with the vertical. If the sphere is initially stationary and the line joining its center to the point at which the string is attached is perpendicular to the vertical, determine the instantaneous $\ddot{\theta}$.



Solutions

1. Circular Hoop*

Since the perpendicular distance of all elements to the axis is r,

$$I = \int r^2 dm = mr^2.$$

2. Spherical Shell*

Define a Cartesian x, y and z coordinate system with the center of the shell as the origin. The moment of inertia about the corresponding axes are

$$I_x = \int (y^2 + z^2) dm,$$

$$I_y = \int (x^2 + z^2) dm,$$

$$I_z = \int (x^2 + y^2) dm.$$

Due to the symmetry of the shell, $I_x = I_y = I_z$. Therefore,

$$\begin{aligned} 3I &= I_x + I_y + I_z \\ &= \int 2(x^2 + y^2 + z^2) dm \\ &= \int 2r^2 dm \\ &= 2mr^2 \\ I &= \frac{2}{3} mr^2. \end{aligned}$$

3. Cube**

(a) Integration

Let the x, y and z-axes be perpendicular to the faces of the cube, with the z-axis being the axis that the moment of inertia is computed with respect to. Define the origin at the center of the cube. The contribution of an infinitesimal box element, $dm = \rho dx dy dz$ where ρ is the mass density, at coordinates

(x, y, z), to the moment of inertia is $(x^2 + y^2)dm$. Therefore,

$$\begin{aligned}
I &= \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{-\frac{l}{2}}^{\frac{l}{2}} (x^2 + y^2)\rho dz dx dy \\
&= \rho \int_{-\frac{l}{2}}^{\frac{l}{2}} \int_{-\frac{l}{2}}^{\frac{l}{2}} l(x^2 + y^2) dx dy \\
&= \rho \int_{-\frac{l}{2}}^{\frac{l}{2}} \left[\frac{lx^3}{3} + ly^2 x \right]_{-\frac{l}{2}}^{\frac{l}{2}} dy \\
&= \rho \int_{-\frac{l}{2}}^{\frac{l}{2}} \left(\frac{l^4}{12} + l^2 y^2 \right) dy \\
&= \rho \left[\frac{l^4}{12} y + \frac{l^2 y^3}{3} \right]_{-\frac{l}{2}}^{\frac{l}{2}} \\
&= \frac{\rho l^5}{6} \\
&= \frac{1}{6} m l^2.
\end{aligned}$$

(b) Squashing and Perpendicular Axis Theorem

To determine the moment of inertia of a cube about the z-axis, squash it along the z-axis to obtain a uniform square. Applying the perpendicular axis theorem to this uniform square,

$$I_z = I_x + I_y,$$

where x and y are axes in the plane of the square through its center of mass, that are parallel to its edges. Due to symmetry, $I_x = I_y$. To compute I_x, we can further squash the square into a uniform rod whose moment of inertia is $\frac{1}{12} ml^2$. Therefore,

$$I_z = 2I_x = \frac{1}{6} ml^2.$$

(c) Scaling Arguments and Parallel Axis Theorem

The moment of inertia of a cube of edge length $2l$ about the z-axis is 32 times that of a cube of edge length l as its mass is eight-fold and its length dimensions are doubled.

$$I_2 = 32I.$$

Furthermore, the former is also composed of the moment of inertia of eight cubes about an axis through one of its edges.

$$I_2 = 8I'.$$

By the parallel axis theorem,

$$I' = I + \frac{1}{2}ml^2.$$

Solving the above equations,

$$I = \frac{1}{6}ml^2.$$

4. Cone**

Let the z-axis be along the height of the cone and define the origin at the vertex of the cone. Let the cone span $z = 0$ to $z = l$ so that the cone is "inverted". Denote ρ as the mass density of the cone. In cylindrical coordinates, the contribution to the moment of inertia — due to an infinitesimal element $dm = \rho r d\phi dr dz$ at z-coordinate z, azimuthal coordinate ϕ and radial coordinate r — is $r^2 dm = \rho r^3 d\phi dr dz$. Hence, we need to integrate this expression over the entire cone. Observe that the limits of integration for r are in fact dependent on z. Explicitly, $r = z\tan\theta$ (by expressing r in terms of z, we must integrate over r before z). Therefore,

$$\begin{aligned}
I &= \int_0^l \int_0^{z\tan\theta} \int_0^{2\pi} \rho r^3 d\phi dr dz \\
&= \rho \int_0^l \int_0^{z\tan\theta} 2\pi r^3 dr dz \\
&= \rho \int_0^l \frac{1}{2}\pi z^4 \tan^4\theta dz \\
&= \frac{1}{10}\rho\pi l^5 \tan^4\theta.
\end{aligned}$$

Since $m = \frac{1}{3}\rho\pi l^3 \tan^2\theta$,

$$I = \frac{3}{10} m l^2 \tan^2\theta.$$

5. Cylinder with Liquid**

The moment of inertia of the shaded part can be computed by subtracting that due to the square from that due to the entire circle and dividing the result by four. The moment of inertia of the circle about the center is

$$\frac{1}{2} m_{circle} R^2 = \frac{1}{2}\sigma\pi R^4.$$

The square has edge length $\sqrt{2}R$. Its moment of inertia about the center is thus

$$\frac{1}{6} m_{square} \cdot 2R^2 = \frac{1}{6} \cdot 2\sigma R^2 \cdot 2R^2 = \frac{2}{3}\sigma R^4.$$

The moment of inertia of the shaded area is thus

$$\frac{\frac{1}{2}\sigma\pi R^4 - \frac{2}{3}\sigma R^4}{4} = \frac{1}{8}\sigma\pi R^4 - \frac{1}{6}\sigma R^4.$$

With regard to the second problem, define the y-axis to be positive downwards and the x-axis to be positive rightwards. We squash the liquid in the z-direction such that its moment of inertia about the axis is identical to that of its cross-section (partially-filled circle) with surface mass density ρl. Now, consider an infinitesimal element $dm = \rho l r d\theta dr$ in polar coordinates about the center of the circle, where r is its radial coordinate and θ is the anti-clockwise vector subtended by its position vector and the positive y-axis. Observe that the limits of integration are coupled in this case. It is more convenient to integrate over θ first before r. For a given radius from the center r, the liquid spans an angle from $-\cos^{-1}\frac{h}{r}$ to $\cos^{-1}\frac{h}{r}$. Meanwhile, r spans from $r = h$ to $r = R$. Since the perpendicular distance between an infinitesimal element and the center is simply r, the moment of inertia of the liquid about the center is

$$
\begin{aligned}
I &= \int_h^R \int_{-\cos^{-1}\frac{h}{r}}^{\cos^{-1}\frac{h}{r}} r^2 \cdot \rho l r d\theta dr \\
&= 2\rho l \int_h^R r^3 \cos^{-1}\frac{h}{r} dr \\
&= \left[\frac{\rho l r^4 \cos^{-1}\frac{h}{r}}{2}\right]_h^R - \int_h^R \frac{\rho l r^4}{2} \cdot -\frac{1}{\sqrt{1-\frac{h^2}{r^2}}} \cdot -\frac{h}{r^2} dr \\
&= \frac{\rho l R^4 \cos^{-1}\frac{h}{R}}{2} - \frac{\rho l}{2}\int_h^R \frac{hr^3}{\sqrt{r^2-h^2}} dr \\
&= \frac{\rho l R^4 \cos^{-1}\frac{h}{R}}{2} - \frac{\rho l}{2}\int_h^R \left(\frac{hr(r^2-h^2)+h^3 r}{\sqrt{r^2-h^2}}\right) dr \\
&= \frac{\rho l R^4 \cos^{-1}\frac{h}{R}}{2} - \frac{\rho l h}{2}\int_h^R r\sqrt{r^2-h^2} dr - \frac{\rho l h^3}{2}\int_h^R \frac{r}{\sqrt{r^2-h^2}} dr,
\end{aligned}
$$

where we have applied integration by parts in the third equality and used the fact that $\frac{d(\cos^{-1} x)}{dx} = -\frac{1}{\sqrt{1-x^2}}$. Adopting the substitutions $u = r^2 - h^2$ and $du = 2rdr$,

$$
\begin{aligned}
I &= \frac{\rho l R^4 \cos^{-1}\frac{h}{R}}{2} - \frac{\rho l h}{4}\int_0^{R^2-h^2} \sqrt{u} du - \frac{\rho l h^3}{4}\int_0^{R^2-h^2} \frac{1}{\sqrt{u}} du \\
&= \frac{1}{2}\rho l R^4 \cos^{-1}\frac{h}{R} - \frac{1}{6}\rho l h (R^2-h^2)^{\frac{3}{2}} - \frac{1}{2}\rho l h^3 \sqrt{R^2-h^2} \\
&= \frac{1}{2}\rho l R^4 \cos^{-1}\frac{h}{R} - \frac{1}{6}\rho l h \sqrt{R^2-h^2}(R^2+2h^2).
\end{aligned}
$$

A simpler way of obtaining this without any integration is as follows. The desired moment of inertia is the moment of inertia of a sector of angle $\theta = 2\cos^{-1}\frac{h}{R}$ minus that of an isosceles triangle, all computed about the same axis stated in the problem.

$$
I = I_{sector} - I_{triangle}.
$$

I_{sector} is simply $\frac{\theta}{2\pi}$ times the moment of inertia of a full circle (of surface mass density ρl obtained after squashing).

$$I_{sector} = \frac{2\cos^{-1}\frac{h}{R}}{2\pi} \cdot \frac{1}{2}\rho l \pi R^2 \cdot R^2 = \frac{1}{2}\rho l R^4 \cos^{-1}\frac{h}{R}.$$

With regard to the moment of inertia of the isosceles triangle, we can divide the isosceles triangle into two right-angled triangles along its perpendicular bisector such that the moment of inertia of the isosceles triangle must be twice the individual moment of inertia of a right-angled triangle, I_{right}, about the proposed axis.

$$I_{triangle} = 2I_{right}.$$

Now, we can apply the parallel axis theorem to conclude that

$$I_{right} = I'_{right} - \left(\frac{1}{2}\rho l h\sqrt{R^2 - h^2}\right)\left(\frac{R^2}{36}\right) + \left(\frac{1}{2}\rho l h\sqrt{R^2 - h^2}\right)\left(\frac{R^2 - h^2}{9} + \frac{4h^2}{9}\right),$$

where I'_{right} is the moment of inertia of a right-angled triangle about a perpendicular axis through the midpoint of its hypotenuse. In writing the above, $\frac{1}{2}\rho l h\sqrt{R^2 - h^2}$ is the mass of the (squashed) right-angled triangle, $\frac{R^2}{36}$ is the squared distance between the midpoint of the hypotenuse and the center of mass and $\frac{R^2-h^2}{9} + \frac{4h^2}{9}$ is the squared distance between the vertex (corresponding to the original axis) and the center of mass. Finally, we know that I'_{right} is simply half the moment of inertia of a uniform rectangle with side lengths $\sqrt{R^2 - h^2}$ and h. Therefore,

$$I'_{right} = \frac{1}{2} \cdot \frac{1}{12}\left(\rho l h\sqrt{R^2 - h^2}h\right)\left(R^2 - h^2 + h^2\right) = \frac{1}{24}\rho l h R^2\sqrt{R^2 - h^2}$$

$$\implies I_{triangle} = \frac{1}{6}\rho l h\sqrt{R^2 - h^2}(R^2 + 2h^2).$$

Substituting this expression back into I,

$$I = I_{sector} - I_{triangle} = \frac{1}{2}\rho l R^4 \cos^{-1}\frac{h}{R} - \frac{1}{6}\rho l h\sqrt{R^2 - h^2}(R^2 + 2h^2).$$

To check our answers, substituting $h = 0$ should yield the moment of inertia of a half-filled circle (which should be $\frac{1}{2} \cdot \frac{1}{2}\rho l \pi R^4 = \frac{1}{4}\rho l \pi R^4$ — half the moment of inertia of a fully-filled circle).

$$I(h=0) = \frac{1}{2}\rho l R^4 \cdot \frac{\pi}{2} - 0 = \frac{1}{4}\rho l \pi R^4.$$

$h = R$ should yield zero as it corresponds to an empty circle.

$$I(h=R) = 0,$$

$h = -R$ should yield the moment of inertia of a full circle — $\frac{1}{2}\rho l \pi R^4$.

$$I(h=-R) = \frac{1}{2}\rho l R^4 \pi - 0 = \frac{1}{2}\rho l \pi R^4.$$

Finally, $h = \frac{R}{\sqrt{2}}$ should yield the answer obtained from the first part of the problem (with $\sigma = \rho l$).

$$I\left(h = \frac{R}{\sqrt{2}}\right) = \frac{1}{2}\rho l R^4 \cdot \frac{\pi}{4} - \frac{1}{6}\rho l \cdot \frac{R}{\sqrt{2}} \cdot \frac{R}{\sqrt{2}} \cdot 2R^2 = \frac{1}{8}\rho l \pi R^4 - \frac{1}{6}\rho l R^4.$$

6. Cube about Any Axis***

Define the x, y and z-axes to be perpendicular to the faces of the cube and the origin at the center of the cube. Now let the direction vector of a general axis A be $\hat{\boldsymbol{d}} = (a, b, c)$ such that $a^2+b^2+c^2 = 1$. Then, the squared perpendicular distance between an infinitesimal box element at position $\boldsymbol{r} = (x, y, z)$ and this axis is

$$\begin{aligned} r_\perp^2 &= |\boldsymbol{r}|^2 - \left(\boldsymbol{r}\cdot\hat{\boldsymbol{d}}\right)^2 \\ &= (1-a^2)x^2 + (1-b^2)y^2 + (1-c^2)z^2 - 2abxy - 2bcyz - 2acxz. \end{aligned}$$

Due to the symmetry of the cube, the terms involving xy, yz and xz will vanish after integrating over the entire distribution.

$$\begin{aligned} I &= \int r_\perp^2 \, dm \\ &= \int \left[(1-a^2)x^2 + (1-b^2)y^2 + (1-c^2)z^2 - 2abxy - 2bcyz - 2acxz\right] dm \\ &= \int \left[(1-a^2)x^2 + (1-b^2)y^2 + (a^2+b^2)z^2\right] dm \\ &= \int \left[a^2(y^2+z^2) + b^2(x^2+z^2) + (1-a^2-b^2)(x^2+y^2)\right] dm \\ &= a^2 I_x + b^2 I_y + (1-a^2-b^2) I_z. \end{aligned}$$

Since $I_x = I_y = I_z = \frac{1}{6}ml^2$,

$$I = I_x = \frac{1}{6}ml^2.$$

7. Cube Fractal***

Consider a fractal that originated from a cube of edge length $3l$. The mass of this fractal is 26 times that of a fractal with length l while its length dimension is 3 times the latter's. Therefore, the moment of inertia of this larger fractal is $26 \times 3^2 = 234$ times that of the original fractal.

$$I_3 = 234I.$$

I_3 is also composed of the moment of inertia of 26 fractals taken about 3 different types of axes. Let the original axis be the z-axis. There are 2 fractals along this z-axis with moment of inertia I. There are 12 fractals about an axis displaced a distance l away from the z-axis, with moment of inertia I_l. There are 12 fractals about an axis displaced a distance $\sqrt{2}l$ away from the z-axis, with moment of inertia $I_{\sqrt{2}l}$. Thus,

$$I_3 = 2I + 12I_l + 12I_{\sqrt{2}l}.$$

By the parallel axis theorem,

$$I_l = I + ml^2,$$

$$I_{\sqrt{2}l} = I + 2ml^2.$$

Solving,

$$I = \frac{9}{52}ml^2.$$

8. Cone about Slanted Axis****

Define the z-axis to be the height of the cone. Then, define axis A to lie strictly in the yz-plane with a direction vector $\hat{\boldsymbol{d}} = (0, \sin\theta, \cos\theta)$, passing through the vertex of the cone. Define the origin at the vertex such that the cone spans from $z = 0$ to $z = -l$. The squared perpendicular distance between an infinitesimal mass element at $\boldsymbol{r} = (x, y, z)$ and axis A is

$$\begin{aligned} r_\perp^2 &= |\boldsymbol{r}|^2 - \left(\boldsymbol{r} \cdot \hat{\boldsymbol{d}}\right)^2 \\ &= x^2 + y^2 \cos^2\theta + z^2 \sin^2\theta - 2yz \sin\theta \cos\theta. \end{aligned}$$

The moment of inertia is then

$$\begin{aligned} I &= \int r_\perp^2 dm \\ &= \int \left(x^2 + y^2 \cos^2\theta + z^2 \sin^2\theta - 2yz \sin\theta \cos\theta\right) dm \\ &= \int \left(x^2 + y^2 \cos^2\theta + z^2 \sin^2\theta\right) dm. \end{aligned}$$

The integral of yz over the mass distribution is again zero due to the fact that for every infinitesimal mass element dm at (y, z), there is a counterpart at $(-y, z)$. However, notice that we still cannot conveniently express the above formula in terms of moment of inertia of the cone about its height $I_z = \int (x^2 + y^2) dm$. That said, we can again leverage on the symmetry of the cone to argue that the moment of inertia stays the same if we swap the y and x-coordinates (we could have defined the x-axis as y and the y-axis as x and the limits of integration over dm would not change).

$$I = \int \left(y^2 + x^2 \cos^2\theta + z^2 \sin^2\theta\right) dm.$$

Summing the two expressions for I,

$$2I = \int \left[(x^2 + y^2)\cos^2\theta + 2z^2 \sin^2\theta\right] dm = I_z \cos^2\theta + \int 2z^2 \sin^2\theta dm.$$

The latter integral can be evaluated easily by slicing the cone into infinitesimal disks of radius $r = -z\tan\theta$ and thickness dz along the z-axis. If we define ρ to be the mass density of the cone,

$$\begin{aligned} \int z^2 dm &= \rho \int_{-l}^{0} z^2 \cdot \pi z^2 \tan^2\theta dz \\ &= \frac{1}{5}\rho\pi l^5 \tan^2\theta \\ &= \frac{3}{5} m l^2 \end{aligned}$$

as the mass of the cone is $m = \rho \cdot \frac{1}{3}\pi l^2 \tan^2\theta \cdot l$. Therefore,

$$\begin{aligned} I &= \frac{I_z \cos^2\theta}{2} + \sin^2\theta \int z^2 dm \\ &= \frac{3}{20} ml^2 \tan^2\theta \cdot \cos^2\theta + \sin^2\theta \cdot \frac{3}{5} ml^2 \\ &= \frac{3}{4} ml^2 \sin^2\theta, \end{aligned}$$

where I_z has been computed in Problem 4 as $\frac{3}{10} ml^2 \tan^2\theta$.

9. Angular Momentum of a Particle*

a) Since $\boldsymbol{L} = 2\alpha t \ln\left(\frac{t}{t+1}\right)\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{k}}$ is constant,

$$\begin{aligned} \frac{d\boldsymbol{L}}{dt} &= 2\alpha \ln\left(\frac{t}{t+1}\right)\hat{\boldsymbol{k}} + 2\alpha t \cdot \frac{t+1}{t} \cdot \frac{1}{(t+1)^2}\hat{\boldsymbol{k}} \\ &= \left[2\alpha \ln\left(\frac{t}{t+1}\right) + \frac{2\alpha}{t+1}\right]\hat{\boldsymbol{k}}. \end{aligned}$$

b) Suppose that the particle indeed reaches $(0, 0, 2)$. Since $\boldsymbol{L} = \boldsymbol{r} \times \boldsymbol{p}$, the angular momentum of the particle must be perpendicular to the position vector of the particle. As the particle is along the z-axis, $L = 0$ which can only occur at $t = 0$. However, because $\boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}$, the torque at $t = 0$ tends to negative infinity at $t = 0$ and points in the z-direction. However, $\boldsymbol{\tau} = \boldsymbol{r} \times \boldsymbol{F}$ which implies that there should be no torque in the z-direction as the particle's position vector is along it — an evident contradiction.

c) The angular momentum of the particle at $t = 1$ is

$$\boldsymbol{L} = -2\alpha \ln 2\hat{\boldsymbol{k}}.$$

Let the momentum of the particle be $\boldsymbol{p} = (p_x, p_y, p_z)$. Then,

$$\boldsymbol{L} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \times \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} 2p_z \\ 0 \\ -2p_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -2\alpha \ln 2 \end{pmatrix}$$

$$\implies p_z = 0 \quad \text{and} \quad p_x = \alpha \ln 2.$$

10. Massive Atwood's Machine*

The moment of inertia of the wheel and axle system is $\frac{1}{2} m_3 r_1^2$ about the center of the wheel. Let the angular velocity of the wheel and axle system be ω, defined to be positive in the anti-clockwise direction. The angular momentum of the entire system (including m_1 and m_2) about the center of the wheel is then

$$L = m_1 r_1^2 \omega - m_2 r_2^2 \omega + \frac{1}{2} m_3 r_1^2 \omega.$$

Note that the speeds of m_1 and m_2 are $r_1\omega$ and $r_2\omega$ by the conservation of string. The rate of change of angular momentum is then

$$\frac{dL}{dt} = \left(m_1 r_1^2 - m_2 r_2^2 + \frac{1}{2} m_3 r_1^2 \right) \alpha.$$

The net torque on this system about the center of the wheel arises from the weights of m_1 and m_2.

$$\tau = m_1 g r_1 - m_2 g r_2.$$

Equating $\tau = \frac{dL}{dt}$,

$$\alpha = \frac{m_1 g r_1 - m_2 g r_2}{m_1 r_1^2 - m_2 r_2^2 + \frac{1}{2} m_3 r_1^2}.$$

11. Yo-yo*

Define clockwise torques to be positive in value and T to be the tension in the string. The only torque on the yo-yo about its center is that due to tension.

$$Tr = \frac{1}{2} m R^2 \alpha.$$

$F = ma$ gives

$$mg - T = ma.$$

For the string to remain taut,

$$a = r\alpha.$$

Solving the above equations,

$$\alpha = \frac{2gr}{R^2 + 2r^2}.$$

12. Rotating Ball**

In the first case, observe that the vertical axis passing through O is the instantaneous axis of rotation of the rotating ball. Therefore, by Eq. (5.2), the angular momentum of the ball with respect to the desired origin (which is an ICoR) is

$$\boldsymbol{L}_1 = I_0 \Omega \hat{\boldsymbol{k}},$$

where $I_0 = \frac{2}{5}mr^2 + mR^2$ is the moment of inertia of the ball with respect to a vertical axis crossing through O, by the parallel axis theorem. Thus,

$$\boldsymbol{L}_1 = \left(\frac{2}{5}mr^2 + mR^2\right)\Omega \hat{\boldsymbol{k}}.$$

Another way of reaching this conclusion is to see that the angular velocity of the ball in the lab frame is $\boldsymbol{\omega} = \boldsymbol{\Omega} + \boldsymbol{0} = \boldsymbol{\Omega}$ by the additive property of angular velocities (where $\boldsymbol{0}$ is the angular velocity of the ball as observed in the rotating frame attached to its center of mass). Applying Eq. (5.5) and using the fact that the center of mass of the ball travels at a speed $v_{CM} = R\Omega$ tangential to its position vector, we obtain the same result. The kinetic energy of the ball is

$$T_1 = \frac{1}{2} I_0 \Omega^2 = \frac{1}{2}\left(\frac{2}{5}mr^2 + mR^2\right)\Omega^2.$$

In the second case, the angular velocity of the ball must be zero in the lab frame for the same side of the ball to be the right-most side constantly. This may seem counter-intuitive at first as we think that the ball is "spinning." However, observe that this implies that the angular velocity of the ball in the rotating frame attached to its center must be $\boldsymbol{\omega}_{rot} = -\boldsymbol{\Omega}$ so that its angular velocity in the lab frame is $\boldsymbol{\Omega} + \boldsymbol{\omega}_{rot} = \boldsymbol{0}$. This must be so in order for the lines joining the center to all points of the ball to not rotate relative to axes fixed in the lab frame — a fact that makes sense. Therefore, the angular momentum and kinetic energy of the ball only stem from the translational component due to the center of mass. Applying Eq. (5.5) yields

$$\boldsymbol{L}_2 = mR^2 \Omega \hat{\boldsymbol{k}}.$$

Applying Eq. (5.6), the kinetic energy of the ball is

$$T_2 = \frac{1}{2} mR^2 \Omega^2.$$

In the third case, observe that the bottom of the ball that is in contact with the table must be an ICoR, as it is not slipping relative to the table that is stationary in the lab frame. In fact, the instantaneous axis of rotation is the line joining O to the bottom of the ball. Since we know that the velocity of the center of the ball is $v_{CM} = R\Omega$ and is directed tangentially in the anti-clockwise direction, the angular velocity of the ball in the lab frame must be

$$\boldsymbol{\omega} = -\frac{R\Omega}{r}\hat{\boldsymbol{R}},$$

where $\hat{\boldsymbol{R}}$ is the unit vector that points radially outwards from the given pivot (point at height r above origin O) to the center of mass of the sphere. This is a consequence of Eq. (3.25). Moving on, by applying $\boldsymbol{L} = \boldsymbol{L}_{CM} + M\boldsymbol{r}_{CM} \times \boldsymbol{v}_{CM}$ with $\boldsymbol{L}_{CM} = I_{CM}\boldsymbol{\omega}$ in this case (due to the symmetry of the ball), the angular momentum of the ball is

$$\boldsymbol{L}_3 = mR^2\Omega\hat{\boldsymbol{k}} - \frac{2}{5}mrR\Omega\hat{\boldsymbol{R}}.$$

Applying Eq. (5.6), the kinetic energy in this case is

$$T_3 = \frac{1}{2} \cdot \frac{2}{5}mr^2 \cdot \frac{R^2\Omega^2}{r^2} + \frac{1}{2}mR^2\Omega^2 = \frac{7}{10}mR^2\Omega^2.$$

The angular momentum of the ball about an origin fixed to its center of mass can be retrieved by subtracting the above expressions for the angular momenta about O by the contribution due to the translational motion of the center of mass ($mR^2\Omega\hat{\boldsymbol{k}}$) about O. Evidently, there is only a rate of change of angular momentum in the third case — signifying that there should be a net external torque about the center of the ball. In fact, we can compute $\frac{d\boldsymbol{L}_{3,CM}}{dt}$ as

$$\frac{d\boldsymbol{L}_{3,CM}}{dt} = -\frac{2}{5}mrR\Omega\frac{d\hat{\boldsymbol{R}}}{dt}.$$

That is, the rate of change only arises from the changing direction of the angular velocity vector of the ball (which is of constant magnitude) — this is no longer a fixed axis rotation but the system is still simple enough to analyze with our tools so far. Observe that since $\hat{\boldsymbol{R}}$ points from the pivot (proposed in the problem) to the center of the ball which rotates at angular velocity $\boldsymbol{\Omega}$, $\hat{\boldsymbol{R}}$ follows suit and rotates at $\boldsymbol{\Omega}$. It was derived in Chapter 3 that the rate of change of a vector of fixed length and rotating at $\boldsymbol{\Omega}$ is

$$\frac{d\hat{\boldsymbol{R}}}{dt} = \boldsymbol{\Omega} \times \hat{\boldsymbol{R}} = \Omega\hat{\boldsymbol{k}} \times \hat{\boldsymbol{R}} = \Omega\hat{\boldsymbol{\theta}},$$

where $\hat{\boldsymbol{\theta}}$ is the unit vector in the tangential direction (positive anti-clockwise). Therefore, the net external torque about the center of the ball must be

$$\sum \boldsymbol{\tau} = \frac{d\boldsymbol{L}_{3,CM}}{dt} = -\frac{2}{5}mrR\Omega^2\hat{\boldsymbol{\theta}}.$$

13. Another Atwood**

Define T to be the tension in the string. Let v_{CM} and a_{CM} be the velocity and the acceleration of the center of mass of the cylinder along the plane, and let a_y be the vertical acceleration of mass m. The positive directions of these quantities are upwards. Finally, let ω and α be the angular velocity and acceleration of the cylinder, positive in the anti-clockwise direction. Applying Newton's second law to the mass,

$$T - mg = ma_y.$$

Next, consider the free-body diagram of the cylinder. To ignore the contact forces (namely friction and the normal force), take the angular momentum about the point of contact between the cylinder and the plane. Then,

$$L = 2mrv_{CM} + I_{CM}\omega$$
$$\frac{dL}{dt} = 2mra_{CM} + mr^2\alpha,$$

as $I_{CM} = mr^2$ for a cylinder. The net torque on the cylinder about the point of contact stems from its weight and tension. Recalling that we defined anti-clockwise to be positive,

$$\tau = T \cdot 2r - 2mgr\sin\theta.$$

Equating $\tau = \frac{dL}{dt}$,

$$2Tr - 2mgr\sin\theta = 2mra_{CM} + mr^2\alpha.$$

For the cylinder to not slip with respect to the plane $a_{CM} = r\alpha$,

$$2Tr - 2mgr\sin\theta = 3mr^2\alpha.$$

Furthermore, by the conservation of string, $a_y = -2r\alpha$. In writing this, we note that the top of the cylinder accelerates tangentially at $2r\alpha$.

$$T - mg = ma_y \implies T - mg = -2mr\alpha.$$

Solving the last two equations,

$$\alpha = \frac{2g}{7r}(1 - \sin\theta).$$

Observe that the judicious choice of locating the origin at the point of contact circumvents the need to consider the static friction acting on the cylinder.

A less direct, but equally valid, solution picks the origin at the center of the cylinder and hence must consider the static friction force f (defined below to be positive downslope but is possibly negative as we may have guessed its direction wrongly). The equation of motion of mass m remains the same.

$$T - mg = ma_y = -2mr\alpha.$$

Applying Newton's second law to the cylinder,

$$T - 2mg\sin\theta - f = 2ma_{CM} = 2mr\alpha.$$

Equating $\tau = I_{CM}\alpha$ about the center of the cylinder,

$$(T + f)r = mr^2\alpha \implies T + f = mr\alpha.$$

Adding the last two equations,

$$2T - 2mg\sin\theta = 3mr\alpha.$$

Subtracting the previous equation by the equation of motion of mass m multiplied by 2,

$$2mg(1 - \sin\theta) = 7mr\alpha$$
$$\alpha = \frac{2g}{7r}(1 - \sin\theta).$$

Notice that we did not have to worry about the actual direction of the static friction force in our calculations. If the final value of f turns out to be negative, it simply means that we made a wrong guess with regard to its direction. In fact, one can substitute α back into the equations above and show that

$$f = -\frac{1}{7}mg - \frac{6}{7}mg\sin\theta,$$

which implies that our hunch was incorrect and that f should really point up-slope.

14. Staying on a Ramp**

If the cylinder does not lose contact with the surface, it will undergo circular motion about the kink for an angle α before sliding down the ramp. Taking the kink as the pivot, the radial component of the cylinder's weight decreases as the angle between its center and the vertical increases. Moreover, the angular velocity of the cylinder increases by the conservation of energy — implying that the required centripetal force increases. Therefore, the normal force, which is directed radially outwards, is minimum when the cylinder has rotated an angle α. We simply have to consider whether it loses contact at this point. The initial kinetic energy of the cylinder is

$$\frac{1}{2}mv_{CM}^2 + \frac{1}{2}I_{CM}\omega^2 = \frac{1}{2}mv_0^2 + \frac{1}{2}\cdot\frac{1}{2}mr^2\omega_0^2 = \frac{3}{4}mv_0^2.$$

Let the angular velocity of the cylinder after it has rotated an angle α be ω. By the conservation of energy,

$$\frac{1}{2}\cdot\frac{3}{2}mr^2\omega^2 = \frac{3}{4}mv_0^2 + mgr(1-\cos\alpha).$$

Note that the expression on the left-hand side describes the rotational energy of the cylinder about the kink (with a moment of inertia about its circumference $\frac{3}{2}mr^2$). In the boundary case where the normal force is zero, the radial component of the cylinder's weight just provides the required centripetal force.

$$mg\cos\alpha = mr\omega^2.$$

Solving the above two equations, the maximum v_0 is thus

$$v_0 = \sqrt{\left(\frac{7}{3}\cos\alpha - \frac{4}{3}\right)gr}.$$

15. Spinning Disc**

Consider the plane of the disc in polar coordinates and define the origin at the center of the disc. Consider an infinitesimal rectangular element at (r,θ) with $dm = \sigma r d\theta dr$. The normal force on this element is

$$dN = \sigma r d\theta dr g.$$

The friction force on this element is $df = \mu\sigma r d\theta dr g$ and the torque due to this friction about the origin is

$$d\tau = -\mu\sigma r^2 d\theta dr g,$$

where the negative sign arises from the fact that the torque's direction is opposite to that of the disc's angular velocity. The total torque on the disc about the origin is then

$$\begin{aligned}
\tau &= \int_0^R \int_0^{2\pi} -\mu\sigma r^2 d\theta dr g \\
&= \int_0^R -2\pi\mu\sigma r^2 g dr \\
&= -\frac{2}{3}\pi\mu\sigma R^3 g \\
&= -\frac{2}{3}\mu m g R \\
\tau &= I_{CM}\alpha, \\
\frac{1}{2}mR^2\alpha &= -\frac{2}{3}\mu m g R \\
\alpha &= -\frac{4\mu g}{3R}, \\
\omega &= \omega_0 - \frac{4\mu g}{3R}t.
\end{aligned}$$

When $\omega = 0$,

$$t = \frac{3R\omega_0}{4\mu g}.$$

16. Spinning Fan**

We will determine the net torque on the entire fan about its central axis by considering that on a single blade. Define the origin to be along the central axis. The drag force on an infinitesimal strip of width a and length dr at a distance r from the central axis is

$$Dv^2 adr = Dr^2\omega^2 adr.$$

The torque about the central axis on this element is thus

$$-Dr^3\omega^2 adr.$$

The total torque on a blade is then

$$\tau_{blade} = -\int_0^b Dr^3\omega^2 adr = -\frac{1}{4}D\omega^2 ab^4.$$

The total torque on the fan about its central axis is then four times that of the above. The moment of inertia of a blade with respect to the central axis is obtained from the parallel axis theorem:

$$I_{blade} = I_{CM} + md^2 = \frac{1}{12}mb^2 + m\frac{b^2}{4} = \frac{1}{3}mb^2.$$

Applying $\tau = I\alpha$ to the entire fan,

$$\frac{4}{3}mb^2\alpha = -D\omega^2 ab^4$$

$$\alpha = -\frac{3Dab^2}{4m}\omega^2.$$

Separating variables and integrating,

$$\int_{\omega_0}^{\omega} \frac{1}{\omega^2} d\omega = \int_0^t -\frac{3Dab^2}{4m} dt$$

$$\frac{1}{\omega_0} - \frac{1}{\omega} = -\frac{3Dab^2}{4m}t$$

$$\omega = \frac{4m\omega_0}{4m + 3Dab^2 t}.$$

17. Sweeping Rod**

Consider an infinitesimal segment of the rod between radial distances r and $r + dr$. After the rod has rotated by an angle θ, this segment would have gathered dust of mass $\sigma r\theta dr$. Therefore, the total moment of inertia due to the mass after an angle θ is

$$I_{dust}(\theta) = \int_0^l \sigma r^3 \theta dr = \frac{\sigma l^4 \theta}{4}.$$

The total angular momentum of the rod and all the dust on the table about the pivot as a function of θ is

$$L = (I_{dust} + I_{rod})\omega.$$

Therefore, the rate of change of angular momentum about the pivot is

$$\frac{dL}{dt} = \frac{dI_{dust}}{dt}\omega = \frac{\sigma l^4 \omega^2}{4},$$

since I_{rod} and ω are constants. This must be equal to the net external torque on the entire system which is exerted along the rod. To maximise the efficacy of a force in generating torque about the pivot, the force must be exerted tangentially at the tip of the rod, a distance l away from the pivot. Then,

$$F = \frac{\sigma l^3 \omega^2}{4}.$$

18. Sphere and Cylinder**

Let ϕ be the angle that the sphere has rotated clockwise in the lab frame (i.e. if we track a point on the sphere that was originally on its top, the angle subtended by the vertical through the center of the sphere and this point will be ϕ). The center of mass of the sphere travels at an azimuthal velocity $(a+b)\dot{\theta}$. The non-slip condition is thus

$$(a+b)\ddot{\theta} = b\ddot{\phi}.$$

Let f and N be the friction and normal forces on the sphere. f is directed tangentially in the anti-clockwise direction as the weight of the sphere tends to pull it down the cylinder, translationally. Applying $F = ma$ to the sphere in the azimuthal direction,

$$mg\sin\theta - f = m(a+b)\ddot{\theta}.$$

Now apply $\tau = I\ddot{\phi}$ to the sphere about its center (note that the angular acceleration of the sphere in the lab frame is truly $\ddot{\phi}$. In the definition of ϕ, we have already included the component of rotation due to increasing θ). The only torque is due to that of friction.

$$fb = \frac{2}{5}mb^2\ddot{\phi}$$

$$f = \frac{2}{5}m(a+b)\ddot{\theta}.$$

Substituting this expression for f into the second equation,

$$mg\sin\theta = \frac{7}{5}m(a+b)\ddot{\theta}$$

$$\ddot{\theta} = \frac{5g\sin\theta}{7(a+b)}.$$

To determine $\dot{\theta}$, use the trick $\ddot{\theta} = \dot{\theta}\frac{d\dot{\theta}}{d\theta}$, separate variables and integrate.

$$\int_0^{\dot{\theta}} \dot{\theta}d\dot{\theta} = \int_0^{\theta} \frac{5g\sin\theta}{7(a+b)}d\theta.$$

After some simplification,

$$\dot{\theta} = \sqrt{\frac{10g}{7(a+b)}(1-\cos\theta)}.$$

When the sphere starts to leave the cylinder, the normal force on the sphere due to the cylinder is zero. Observe that $mg\cos\theta - N$ provides the centripetal force which is of magnitude $m(a+b)\dot{\theta}^2$. Therefore, the sphere loses contact with the cylinder when

$$mg\cos\theta = m(a+b)\dot{\theta}^2.$$

Substituting the expression for $\dot{\theta}$ and simplifying would yield

$$\theta_0 = \cos^{-1}\frac{10}{17}.$$

19. Sphere on Block**

Let a_{CM}, α and a_{block} denote the acceleration of the sphere, angular acceleration of the sphere (positive anti-clockwise) and the acceleration of the block respectively. Let f be the friction on the sphere due to the block (defined to be positive rightwards). Then,

$$F - f = Ma_{block}$$
$$f = ma_{CM}$$
$$fr = \frac{2}{5}mr^2\alpha.$$

For the sphere to not slip relative to the block,

$$a_{CM} - r\alpha = a_{block}.$$

Solving the above equations,

$$f = \frac{F}{1 - \frac{3M}{2m}}.$$

The normal force between them is mg. Thus, from $\left|\frac{f}{N}\right| \leq \mu$, the maximum F is governed by

$$\left|\frac{F}{mg - \frac{3}{2}Mg}\right| \leq \mu$$
$$F \leq \mu\left|mg - \frac{3}{2}Mg\right|.$$

To determine the displacement of the center of the sphere, observe the sign of f; if $m > \frac{3}{2}M$, the center of the sphere will move forwards. Otherwise, if $m < \frac{3}{2}M$, the center of the sphere will move backwards.

20. Pendulum***

Define θ and ϕ as the angular coordinates depicted in the figure below.

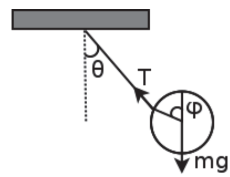


Figure 5.18: θ and ϕ

Define the horizontal x-axis and the vertical y-axis to be positive rightwards and upwards respectively. Then, the coordinates of the center of mass of the sphere are

$$x_{CM} = l\sin\theta + r\sin\phi,$$
$$y_{CM} = -l\cos\theta - r\cos\phi.$$

Therefore,

$$\ddot{x}_{CM} = -l\sin\theta\dot{\theta}^2 + l\cos\theta\ddot{\theta} - r\sin\phi\dot{\phi}^2 + r\cos\phi\ddot{\phi},$$
$$\ddot{y}_{CM} = l\cos\theta\dot{\theta}^2 + l\sin\theta\ddot{\theta} + r\cos\phi\dot{\phi}^2 + r\sin\phi\ddot{\phi}.$$

By considering forces on the sphere,

$$-T\sin\theta = m\ddot{x}_{CM},$$
$$T\cos\theta - mg = m\ddot{y}_{CM}.$$

The tension T makes an angle $\phi - \theta$ clockwise from the radial direction (extending from the center of the sphere). Therefore, the torque about the center of mass of the sphere, positive anti-clockwise, is

$$\tau = -Tr\sin(\phi - \theta) = Tr\sin(\theta - \phi).$$

Equating this with the rate of change of angular momentum of the sphere with respect to the center of mass, we have

$$\frac{2}{5}mr^2\ddot{\phi} = Tr\sin(\theta - \phi).$$

Furthermore, if we substitute the initial conditions $\theta = \theta_0$, $\dot{\theta} = 0$, $\dot{\phi} = 0$ and $\phi = \frac{\pi}{2}$ radians into the above equations,

$$-T \sin\theta_0 = ml \cos\theta_0 \ddot{\theta},$$

$$T \cos\theta_0 - mg = ml \sin\theta_0 \ddot{\theta} + mr\ddot{\phi},$$

$$\frac{2}{5} mr^2 \ddot{\phi} = -Tr \cos\theta_0.$$

Solving the above equations simultaneously, the instantaneous $\ddot{\theta}$ is

$$\ddot{\theta} = -\frac{2g \sin\theta_0}{(2 + 5\cos^2\theta_0)l}.$$

Chapter 6

Energy and Momentum

This chapter will introduce the conservation of linear momentum, angular momentum and energy, which are powerful alternatives to the dynamical laws. Though they are derived entirely from Newton's laws in this chapter, these conservation laws are principles — a creed — that we generally believe in and abide by, even beyond classical mechanics. The earlier sections will be heavy on derivations and can be skipped if the reader is more interested in the application of these concepts.

6.1 Linear Momentum

Conservation of Linear Momentum

Recall that the net external force on a system of particles is directly proportional to the rate of change of its total momentum

$$\sum \boldsymbol{F} = \frac{d\boldsymbol{p}}{dt}.$$

If the net external force on a system is zero,

$$\frac{d\boldsymbol{p}}{dt} = 0$$
$$\boldsymbol{p} = \boldsymbol{p}_0$$

for some constant momentum $\boldsymbol{p}_0$. This is the law of conservation of linear momentum, which states that the total linear momentum of a system is conserved if no net external force acts on it.

Impulse-Momentum Theorem

An alternate perspective of the second law is obtained from separating variables and integrating.

$$\int_{t_0}^{t_1} \left(\sum \boldsymbol{F}\right) dt = \int_{\boldsymbol{p}_0}^{\boldsymbol{p}_1} d\boldsymbol{p}.$$

If we define the impulse delivered by an external force $\boldsymbol{F}_i$ during a time interval between t_0 and t_1 to be

$$\boldsymbol{J} = \int_{t_0}^{t_1} \boldsymbol{F}_i \, dt, \tag{6.1}$$

the first equation can be rewritten as

$$\sum \boldsymbol{J} = \Delta \boldsymbol{p}. \tag{6.2}$$

This is known as the impulse-momentum theorem which states that the change in the linear momentum of a system during a time interval equals the total impulse delivered to the system during that interval.

This new formulation is particularly enlightening in the case of impulsive forces — such as the normal forces between colliding particles — which are gargantuan in magnitude but act over a small time interval. The "role" of such forces is then to impart momentum into systems over a short time interval.

6.2 Angular Momentum

Conservation of Angular Momentum

The relationship between the net external torque and the rate of change of angular momentum of a system with respect to a non-accelerating pivot is

$$\sum \boldsymbol{\tau} = \frac{d\boldsymbol{L}}{dt}.$$

If the net external torque is zero, the angular momentum of a system about the same pivot is conserved.

$$\boldsymbol{L} = \boldsymbol{L}_0.$$

For a single particle, the conservation of angular momentum can also be derived from analyzing its equation of motion in cylindrical coordinates.

The azimuthal equation reads

$$F_\theta = ma_\theta = m(r\ddot{\theta} + 2\dot{r}\dot{\theta}),$$

where r is the perpendicular distance of the particle from the z-axis. If the particle experiences no azimuthal force (and thus no torque in the z-direction),

$$m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = \frac{m(r^2\ddot{\theta} + 2r\dot{r}\dot{\theta})}{r} = \frac{m}{r} \cdot \frac{d(r^2\dot{\theta})}{dt} = 0$$
$$\implies mr^2\dot{\theta} = L_0$$

for some constant L_0. That is, the z-component of the angular momentum of the particle is conserved. A paramount special case of the conservation of angular momentum pertains to the motion of a particle under the influence of a strictly radial force. That is, the force on the particle is always parallel to its position vector with respect to a certain origin.

$$\boldsymbol{F} = F\hat{\boldsymbol{r}}.$$

The net torque on the particle about the origin is then

$$\boldsymbol{\tau} = \boldsymbol{r} \times F\hat{\boldsymbol{r}} = 0.$$

Therefore, the angular momentum of the particle with respect to that origin is conserved.

Problem: A mass m is initially undergoing uniform circular motion on a frictionless, horizontal table at a radius r_0 and tangential velocity v_0 — the centripetal force is provided by the tension that you exert on the string. Suppose that you start to pull the string slowly such that the radial distance r slowly decreases, determine the tangential velocity of the mass $v(r)$.

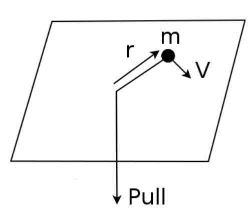


Figure 6.1: Pulling a mass

The tension is solely in the radial direction relative to an origin defined at the hole. Thus, the angular momentum of the block relative to the origin is conserved.

$$mr_0v_0 = mrv$$
$$v = \frac{r_0v_0}{r}.$$

Problem: A uniform sphere of mass m, radius r and initial angular velocity ω_0 rolls with slipping on a rough, horizontal ground. If the initial velocity of its center of mass is zero, determine the final velocity v of the center of mass after a long time.

The friction f on the circumference of the sphere accelerates it translationally and decreases its angular velocity. We can in fact solve this problem using methods established so far. Applying $F = ma$ and $\tau = I\alpha$ about the center,

$$f = ma = m\frac{dv}{dt},$$
$$-fr = I\alpha = \frac{2}{5}mr^2\frac{d\omega}{dt},$$

where the negative sign indicates that the torque due to friction tends to oppose the angular velocity of the sphere. Dividing the second equation by the first and rearranging,

$$\int_0^v dv = -\frac{2}{5}r\int_{\omega_0}^{\omega} d\omega$$
$$v = \frac{2}{5}r\omega_0 - \frac{2}{5}r\omega.$$

The sphere ceases to slip with respect to the ground when $v = r\omega$ such that v at this juncture is

$$v = \frac{2}{5}r\omega_0 - \frac{2}{5}v$$
$$v = \frac{2}{7}r\omega_0.$$

Instead of analyzing the problem in terms of forces and torques, we can also calculate the angular momentum of the sphere with respect to a pivot on the ground. Since the line of action of the friction force passes through this pivot (enabling us to bypass friction entirely) and because the torques due to the vertical forces on the sphere are balanced, the total angular momentum of the sphere with respect to this pivot must be conserved. By the conservation of angular momentum about the pivot on the ground,

$$\frac{2}{5}mr^2\omega_0 = \frac{2}{5}mr^2\omega + mrv = \frac{7}{5}mrv$$
$$v = \frac{2}{7}r\omega_0.$$

Angular Impulse-Momentum Theorem

Similarly, we can define the angular impulse delivered by an external torque $\boldsymbol{\tau}$ about an origin to be

$$\boldsymbol{I} = \int_{t_0}^{t_1} \boldsymbol{\tau} dt \tag{6.3}$$

over a time interval from t_0 to t_1. Then,

$$\sum \boldsymbol{I} = \Delta \boldsymbol{L}. \tag{6.4}$$

This is the angular impulse-momentum theorem. Now, consider the case where there is only a single impulsive force $\boldsymbol{F}(t)$ which acts at a constant position $\boldsymbol{R}$ on an object over a short period of time. Then,

$$\boldsymbol{I} = \boldsymbol{R} \times \int \boldsymbol{F}(t) dt = \boldsymbol{R} \times \boldsymbol{J}$$
$$\implies \Delta \boldsymbol{L} = \boldsymbol{R} \times \Delta \boldsymbol{p}. \tag{6.5}$$

Problem: A cylinder of radius r initially rolls without slipping on a horizontal ground at angular velocity ω. Determine the vertical height h above the ground that a horizontal pole should be used to hit the cylinder such that it immediately rolls without slipping in the opposite direction.

Without loss of generality, assume that the cylinder initially rolls towards the right (positive x-direction) and possesses clockwise angular velocity ω. Let the change in the linear momentum of the cylinder be Δp. Since this must be a negative value for the cylinder to roll backwards, the cylinder can only be hit from the right. Furthermore, by applying Eq. (6.5), the change in its angular momentum about its center is $(h - r)\Delta p$ (defined to be positive clockwise). Thus, the final momentum and angular momentum of the cylinder after the collision are

$$p' = mv' = mr\omega + \Delta p,$$
$$L' = \frac{1}{2}mr^2\omega' = \frac{1}{2}mr^2\omega + (h-r)\Delta p,$$

where v' is the final velocity of the center of the cylinder and ω' is its final angular velocity. The non-slip condition after the collision is

$$v' = r\omega'.$$

Multiplying the first equation by $\frac{1}{2}r$ and subtracting the second,

$$h = \frac{3}{2}r.$$

6.3 Work and Energy

6.3.1 *Work*

If a particle experiences an infinitesimal displacement $d\boldsymbol{r}$ under the influence of a instantaneous force $\boldsymbol{F}$, the infinitesimal work done on the particle due to the force is defined to be

$$dW = \boldsymbol{F} \cdot d\boldsymbol{r}.$$

The total work done on a particle by a force $\boldsymbol{F}$ — which is a general function of time, position and other variables — as it travels along a path P from position vectors $\boldsymbol{r}_0$ to $\boldsymbol{r}_1$ is

$$W = \int_{\boldsymbol{r}_0}^{\boldsymbol{r}_1} \boldsymbol{F} \cdot d\boldsymbol{r}, \tag{6.6}$$

where $\boldsymbol{r}$ is the instantaneous position vector of the particle and $d\boldsymbol{r}$ is the infinitesimal displacement of the particle along path P. The above is an example of a line integral which can be illustrated by Fig. 6.2 below.

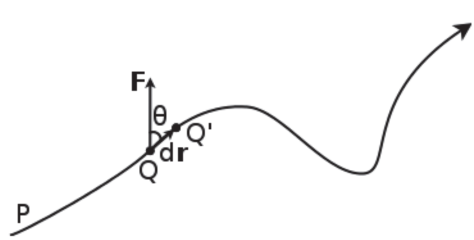


Figure 6.2: Line integral

Let Q be a point on the path P, at which a force $\boldsymbol{F}$ acts ($\boldsymbol{F}$ could be any function along the path P). Let Q' be an adjacent point on P after an infinitesimal displacement $d\boldsymbol{r}$. Then, the work done by the force $\boldsymbol{F}$ as the particle travels from Q to Q' is

$$dW = \boldsymbol{F} \cdot d\boldsymbol{r} = F|d\boldsymbol{r}|\cos\theta,$$

where θ is the instantaneous angle between the instantaneous $\boldsymbol{F}$ and instantaneous $d\boldsymbol{r}$. We then repeat this for all adjacent points along P and sum them up — leading to the integral in Eq. (6.6).

In general, we would need to know the function $\boldsymbol{F}$ and the evolution of the particle's position to determine W. In the special case where $\boldsymbol{F}$ is a constant, the integral can be evaluated trivially.

$$W = \boldsymbol{F} \cdot \int_{\boldsymbol{r}_0}^{\boldsymbol{r}_1} d\boldsymbol{r} = \boldsymbol{F} \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_0).$$

For example, if $\boldsymbol{F}$ is the weight of a particle of mass m, $\boldsymbol{F} = (0, 0, -mg)$,

$$W = \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta z \end{pmatrix} = -mg\Delta z,$$

where Δz is the change in the z-coordinate of the particle. Finally, the work done on a particle can also be expressed as

$$\begin{aligned} W &= \int_{\boldsymbol{r}_0}^{\boldsymbol{r}_1} \boldsymbol{F} \cdot \frac{d\boldsymbol{r}}{dt} dt \\ &= \int_{\boldsymbol{r}_0}^{\boldsymbol{r}_1} \boldsymbol{F} \cdot \boldsymbol{v}\, dt. \end{aligned}$$

The significance of the integrand will be evident soon.

Work Done on a Rigid Body

For a particle on a rigid body, its instantaneous velocity can be expressed as

$$\boldsymbol{v} = \boldsymbol{v}_{CM} + \boldsymbol{\omega} \times \boldsymbol{r}',$$

where $\boldsymbol{v}_{CM}$ is its velocity of its center of mass and $\boldsymbol{r}'$ is the vector pointing from the center of mass to that particle. Therefore, the work done on the rigid body by a force $\boldsymbol{F}$ at a point corresponding to position $\boldsymbol{r}'$ over an infinitesimal time interval is

$$\begin{aligned} dW &= \boldsymbol{F} \cdot \boldsymbol{v} dt \\ &= \boldsymbol{F} \cdot \boldsymbol{v}_{CM} dt + \boldsymbol{F} \cdot (\boldsymbol{\omega} \times \boldsymbol{r}') dt. \end{aligned}$$

The second term can be simplified via the scalar product identity.

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b}),$$

such that $\boldsymbol{F} \cdot (\boldsymbol{\omega} \times \boldsymbol{r}') = \boldsymbol{\omega} \cdot (\boldsymbol{r}' \times \boldsymbol{F}) = \boldsymbol{\omega} \cdot \boldsymbol{\tau}$ where $\boldsymbol{\tau}$ is the instantaneous torque on the rigid body about its center of mass. Thus,

$$\begin{aligned} dW &= \boldsymbol{F} \cdot \boldsymbol{v}_{CM} dt + \boldsymbol{\tau} \cdot \boldsymbol{\omega} dt \\ &= \boldsymbol{F} \cdot d\boldsymbol{r}_{CM} + \boldsymbol{\tau} \cdot d\boldsymbol{\theta}, \end{aligned}$$

where $d\boldsymbol{r}_{CM}$ is the infinitesimal displacement of the center of mass and $d\boldsymbol{\theta}$ is the infinitesimal rotation of the rigid body about the center of mass. Thus, the total work done by a force on a rigid body, under the assumption of a fixed axis rotation about the z-axis, is

$$W = \int_{\boldsymbol{r}^0_{CM}}^{\boldsymbol{r}^1_{CM}} \boldsymbol{F} \cdot d\boldsymbol{r}_{CM} + \int_{\theta_0}^{\theta_1} \tau_z d\theta, \tag{6.7}$$

where $\boldsymbol{r}^0_{CM}$ and $\boldsymbol{r}^1_{CM}$ are the initial and final positions of the center of mass respectively while θ_0 and θ_1 are the initial and final angles that the rigid body has rotated about its center of mass.[1]

Problem: A sphere of mass m, radius r and initial angular velocity ω_0 rolls with slipping on a rough ground. If the initial velocity of its center of mass is zero, determine the total work done by friction on the sphere.

Observing that the net work on the sphere is that due to friction, let the coefficient of kinetic friction between the sphere and the ground be μ. As the force and torque about the center of mass due to friction are constant, we simply have to evaluate the linear displacement of the center of mass s and the angle that the body has rotated until it stops slipping, θ. The net force, which is friction, causes the sphere to accelerate at $a = \mu g$ while the torque due to friction about its center of mass causes its angular acceleration to be $\alpha = -\frac{5\mu g}{2r}$ (negative as it opposes ω_0). Recall that in the previous question,

[1]Note that in defining these two angles, we have assumed the axis of rotation to be fixed as it is impossible to represent a general rotation via a scalar or vector.

we determined that the final velocity of the center of mass is $\frac{2r\omega_0}{7}$. Thus, the time taken for the sphere to stop slipping is

$$t = \frac{2r\omega_0}{7\mu g}.$$

From the kinematics equations,

$$s = \frac{1}{2}at^2 = \frac{1}{2}\mu g\left(\frac{2r\omega_0}{7\mu g}\right)^2 = \frac{2r^2\omega_0^2}{49\mu g},$$

$$\theta = \omega_0 t + \frac{1}{2}\alpha t^2 = \frac{2r\omega_0^2}{7\mu g} - \frac{1}{2}\cdot\frac{5\mu g}{2r}\cdot\left(\frac{2r\omega_0}{7\mu g}\right)^2 = \frac{9r\omega_0^2}{49\mu g}.$$

The total work done by friction is therefore

$$\begin{aligned} W &= F\cdot s + \tau_z\cdot\theta \\ &= \mu mg\cdot\frac{2r^2\omega_0^2}{49\mu g} - \mu mgr\cdot\frac{9r\omega_0^2}{49\mu g} \\ &= -\frac{1}{7}mr^2\omega_0^2. \end{aligned}$$

Do not sweat over the significance of the negative sign. The negative sign of W has nothing to do with direction; $\frac{1}{7}mr^2\omega_0^2$ is not the net work done to the left. Rather, $W = -\frac{1}{7}mr^2\omega_0^2$, as a whole, is simply the work done. A direction cannot be assigned to a scalar. There is also no point in describing the magnitude of W (which, by the way, is still $-\frac{1}{7}mr^2\omega_0^2$ and not a positive value[2]) as a scalar only has magnitude and is the magnitude in the first place.

6.3.2 *Work-Energy Theorem*

The work-energy theorem states that the total work done by external forces W on a particle or a rigid body, undergoing a fixed axis rotation, is equal to its change in kinetic energy, ΔT.

$$W = \Delta T. \tag{6.8}$$

Proof: We shall drop the limits lest the expressions become too cluttered. Just remember that the limits describe the relevant quantities at the initial

[2] Do not confuse the magnitude of a vector $|\boldsymbol{A}|$ with taking the absolute value of a scalar.

and final states. For a particle,

$$\begin{aligned}W &= \int \left(\sum \boldsymbol{F}\right) \cdot \boldsymbol{v} dt \\ &= \int m\boldsymbol{v} \cdot \frac{d\boldsymbol{v}}{dt} dt \\ &= \int m\boldsymbol{v} \cdot d\boldsymbol{v} \\ &= \int \frac{1}{2} m \, d(\boldsymbol{v} \cdot \boldsymbol{v}) \\ &= \Delta \left(\frac{1}{2} m v^2\right),\end{aligned}$$

where $\frac{1}{2}mv^2$ is defined as the kinetic energy of a particle of mass m and speed v. Similarly, for a rigid body under a fixed axis rotation,

$$\begin{aligned}W &= \int \left(\sum \boldsymbol{F}\right) \cdot \boldsymbol{v}_{CM} dt + \int \tau_z \omega dt \\ &= \int M\boldsymbol{v}_{CM} \cdot \frac{d\boldsymbol{v}_{CM}}{dt} dt + \int I_{CM}\omega \cdot \frac{d\omega}{dt} dt \\ &= \int \frac{1}{2} M d(\boldsymbol{v}_{CM} \cdot \boldsymbol{v}_{CM}) + \int I_{CM}\omega \, d\omega \\ &= \Delta \left(\frac{1}{2} M v_{CM}^2 + \frac{1}{2} I_{CM}\omega^2\right).\end{aligned}$$

We emphasize that τ_z is computed with respect to the center of mass such that $\tau_z = \frac{dL_z}{dt} = \frac{d(I_{CM}\omega)}{dt} = I_{CM}\frac{d\omega}{dt}$. Recall that we have derived $\frac{1}{2}Mv_{CM}^2 + \frac{1}{2}I_{CM}\omega^2$ in Chapter 5 from summing the kinetic energies of individual mass elements. Therefore, the expression in the brackets represents the total kinetic energy of the rigid body. Finally, recall that the internal forces in a rigid body result in no net force or torque — implying that W is simply the total work done by external forces.

Problem: Verify that the total work done by friction on the sphere, in the previous question, is indeed $-\frac{1}{7}mr^2\omega_0^2$ by computing its change in kinetic energy.

The initial kinetic energy is

$$T_0 = \frac{1}{2} I_{CM}\omega_0^2 = \frac{1}{5} m r^2 \omega_0^2.$$

The final kinetic energy, with $I_{CM} = \frac{2}{5}mr^2$ for a sphere, is

$$\begin{aligned} T_1 &= \frac{1}{2}mv^2 + \frac{1}{2}I_{CM}\omega^2 \\ &= \frac{1}{2}m\left(\frac{2}{7}r\omega_0\right)^2 + \frac{1}{2}\cdot\frac{2}{5}mr^2\cdot\left(\frac{2}{7}\omega_0\right)^2 \\ &= \frac{2}{35}mr^2\omega_0^2. \end{aligned}$$

By the work-energy theorem, the work done by friction is

$$W = \Delta T = -\frac{1}{7}mr^2\omega_0^2.$$

Power

Considering the equivalence of work and the change in kinetic energy of a particle or rigid body, the instantaneous power delivered to a particle or rigid body by a force $\boldsymbol{F}$ is defined to be the rate of work done by that force.

$$P = \frac{dW}{dt} = \boldsymbol{F}\cdot\boldsymbol{v},$$

where $\boldsymbol{v}$ is the velocity of the point of application.

Problem: If the drag force in air is proportional to the squared speed of an object and P_0 is the power required to push a cart on a frictionless, horizontal ground at a constant horizontal speed v, determine the power required to push a cart at $2v$.

Since the drag force is proportional to the squared speed of the cart, the power delivered by the drag force, and hence the power required, is proportional to the cubed speed. Thus, $8P_0$ is necessary.

6.3.3 *Conservation of Energy*

Energy of a Particle or Rigid Body

Suppose that we are able to define a function $U(\boldsymbol{r})$ for a force $\boldsymbol{F}$ on a particle or rigid body[3] such that

$$U(\boldsymbol{r}) = -\int_{\boldsymbol{r}_0}^{\boldsymbol{r}} \mathbf{F}\cdot d\boldsymbol{r} = -W. \tag{6.9}$$

[3] U would generally depend on the position vectors of all particles on a rigid body.

Then, assuming that $\boldsymbol{F}$ is the only external force present, the work-energy theorem implies that

$$\begin{aligned} U(\boldsymbol{r}) &= T_0 - T \\ T + U(\boldsymbol{r}) &= E \end{aligned} \tag{6.10}$$

for some constant E, which has been substituted for T_0 — the kinetic energy of the particle or rigid body at the state corresponding to $\boldsymbol{r_0}$. Forces for which such a function $U(\boldsymbol{r})$ can be defined are known as conservative forces, and $U(\boldsymbol{r})$ is known as the potential energy associated with the force $\boldsymbol{F}$. Furthermore, the quantity E refers to the total mechanical energy of the object. If there are multiple conservative forces acting on a particle or rigid body, its total mechanical energy is simply the sum of the individual potential energies and its kinetic energy.

The above equation describes the prevalent law of conservation of energy for a single particle or rigid body. In the absence of a net non-conservative force whose work performed cannot be represented by a potential energy function, the total mechanical energy of a particle or rigid body is conserved. This provides us with a powerful alternative to analyzing forces and torques — energy. From the above derivation, we also see that work and potential energy are essentially two different ways of expressing the same idea if we are able to define a potential energy function $U(\boldsymbol{r})$.

The potential energy is essentially a book-keeping device and is not a tangible form of energy that an object physically possesses. Instead, we are associating the object with the potential amount of work that can be performed by conservative forces in defining its potential energy — analogous to writing an additional amount of money on your checkbook in anticipation of what others owe you. When work is finally performed by the conservative forces, the kinetic energy of the body increases but its associated potential energy decreases — akin to how you physically possess more money when others pay you back but they owe you less. There is an interconversion of currency and debit, similar to that between kinetic and potential energies, but the total value (mechanical energy) remains constant. Finally, since the net external force required to balance the conservative force is $-\boldsymbol{F}$, the potential energy at a point is also the work done by an external force in bringing a particle from $\boldsymbol{r_0}$ to $\boldsymbol{r}$ without a change in kinetic energy — this is the loan that you gave out and thus, what others owe you. So under what conditions can we define such a potential energy function?

Firstly, by observing $U(\boldsymbol{r})$, we see that it is strictly dependent on the position of the particle. Thus, our force F must also strictly be a sole function of position — it cannot be a function of velocity, for instance. Note that though the magnitude of kinetic friction is constant, its direction is opposite to the relative velocity between the particle and a surface — causing it to implicitly depend on velocity. Thus, friction is a non-conservative force.

Next, the work done by the force in bringing a particle from $\boldsymbol{r}_0$ to $\boldsymbol{r}_1$ must be path-independent. This is due to the sole dependence of the left-hand side of Eq. (6.9) on the current position of the particle and independence of how it got there. Another corollary of this is that the work done by a force in bringing a particle around a loop and back to its initial position must be zero. This is evidently contravened by friction, which constantly performs negative work on a particle. As an aside, the mathematical criteria necessary for the fulfilment of this condition are: a vector known as the curl of the force must be null everywhere, and the force must not have any singularities (discontinuities). The curl in Cartesian coordinates of a vector $\boldsymbol{F}$ is written as

$$\nabla \times \boldsymbol{F} = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \times \begin{pmatrix} F_x \\ F_y \\ F_z \end{pmatrix} = \begin{pmatrix} \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \\ \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \\ \frac{\partial F_y}{\partial x} - \frac{\partial F_x}{\partial y} \end{pmatrix}. \tag{6.11}$$

We shall state this condition without proof as it is not of particular interest to us.

From the expression for potential energy, we can actually write a conservative force $\boldsymbol{F}$ in Cartesian coordinates, as

$$\boldsymbol{F} = -\frac{\partial U}{\partial x}\hat{\boldsymbol{i}} - \frac{\partial U}{\partial y}\hat{\boldsymbol{j}} - \frac{\partial U}{\partial z}\hat{\boldsymbol{k}}. \tag{6.12}$$

Thus, a conservative force is equal to the negative of its potential energy gradient.

Examples of Potential Energies

Observe that the definition of a potential energy entails a reference point $\boldsymbol{r}_0$ at which the potential energy is set to be zero — that is, an absolute potential energy does not exist. For instance, it is meaningless to define the gravitational potential energy of an object as mgh. We must have a point of reference — for example, h could have been measured relative to the ground at which the gravitational potential energy has been taken to be zero.

Gravitational Potential Energy

If the z-axis is positive vertically upwards and if the gravitational field strength g is constant throughout all space and time, the gravitational force on a particle or rigid body is

$$\boldsymbol{F} = -mg\hat{\boldsymbol{k}}$$

everywhere, and at every instant. The gravitational potential energy is then

$$U_G = -\int_{h_0}^{h_1} -mg\hat{\boldsymbol{k}} \cdot d\boldsymbol{r} = mg(h_1 - h_0),$$

where h_1 is the current z-coordinate of the particle and h_0 is a reference z-coordinate. If h_0 is defined to be at the origin,

$$U_G = mgh_1.$$

Potential Energy of a Spring with a Fixed End

Consider a spring with one end fixed to some entity, such as a wall. Define the origin at the fixed end and attach a particle or rigid body to the other end. Then the force on the object due to the spring is

$$\boldsymbol{F} = -k(r - l)\hat{\boldsymbol{r}},$$

where $\boldsymbol{r}$ is the position vector of the free end of the spring and l is the length of its relaxed state. Since we have assumed the work done by the spring to be independent of the process in defining a potential energy function, consider the simplest case, where the spring is aligned along the x-axis and subsequently stretched or compressed along it. Then,

$$F = -k(x - l).$$

Defining the reference point to be at $x = l$ (relaxed state), the potential energy of the spring is then

$$U_S = -\int_{l}^{x} -k(x - l)dx = \frac{1}{2}k(x - l)^2.$$

Potential Energy Diagrams

As an aside, it is useful to plot a graph of potential energy $U(\boldsymbol{r})$ against $\boldsymbol{r}$ when the total mechanical energy of our system is conserved. We can then draw a horizontal line $y = E$, where E is the total mechanical energy of our system which constrains the motion of our system. The particle or system can only exist in regions where the potential energy is smaller or equal to the total mechanical energy (i.e. below or on the horizontal line) because its kinetic energy must be non-negative. Furthermore, the particle changes its direction of motion at the points of intersection of the potential energy curve and the horizontal line as it stops at the boundary (kinetic energy is zero) and must thus turn due to the conservative force acting on it.

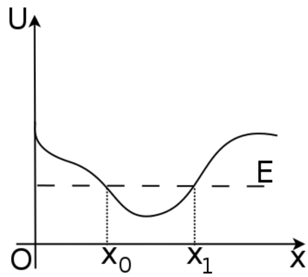


Figure 6.3: Potential energy against position

Consider the above potential energy diagram of a particle moving in a single direction. The particle can only move within the region confined by x_0 and x_1 as it does not have sufficient energy to overcome the potential energy barrier. At the extrema located at x_0 and x_1, the particle will have zero kinetic energy and thus, zero velocity. Since force is equal to the negative of the potential gradient (i.e. acts in the direction opposite to the potential gradient), when the particle is at x_0 and x_1, a force will act to bring the particle towards a lower potential energy — thus imposing restrictions on the region that it can move in. An apt analogy would be a ball on the bottom of a hill. If its total energy is insufficient, it will never be able to roll up the hill, at least in classical mechanics.

Note that potential energy diagrams can also be useful in determining whether an object, under the sole influence of conservative forces, is in a stable or unstable equilibrium. An object will be at equilibrium at an extremum of the potential energy graph as $F = -\frac{dU}{dr} = 0$. Generally, if an object is at a maximum, it will be in an unstable equilibrium as any deviation will lead to a force which tends to thrust the object away from the equilibrium position. On the other hand, if an object is at a minimum, any small deviation could be corrected by a force that is directed towards the minimum.

Finally, let us try to tie together the concepts that we have highlighted so far by analyzing falling dominos!

Problem: Model an array of dominoes as a line of vertical sticks of length l that are separated by a horizontal distance $d < l$ on a rough ground with a coefficient of static friction μ. Suppose we give the first stick a push such that it topples, while its end on the ground remains stationary. It collides with the next stick — causing the next stick to subsequently topple and collide with its neighboring stick and so on. Assuming that only one collision occurs between successive sticks and that the ith stick does not slip relative to the ground until it collides with the $(i+1)$th stick, determine the condition for the $(i+1)$th stick to immediately rotate without slipping relative to the ground after the ith stick collides with it.

Assuming that the previous condition is fulfilled and that energy is conserved across all collisions, show that it is possible to give the first stick a certain initial angular velocity such that all subsequent sticks possess a common angular velocity ω_f, immediately after they have been knocked by their predecessors. Determine ω_f.

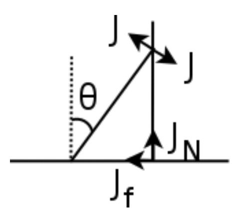


Figure 6.4: Collision between ith and $(i+1)$th sticks

Consider the collision between the ith and $(i+1)$th sticks depicted in Fig. 6.4. Suppose that the impulse delivered between the sticks during the collision is J. In addition to this impulse, the $(i+1)$th stick experiences impulses J_f and J_N due to friction and the normal force from the ground too. In order for the $(i+1)$th stick to not slip with respect to the ground, it must rotate about its stationary end on the ground immediately after the collision. This firstly requires the vertical velocity of its center of mass to be zero. That is,

$$J_N = J\sin\theta,$$

where $\sin\theta = \frac{d}{l}$ (we shall use θ instead of d and l in this problem to simplify our variables). Furthermore, the instantaneous center of rotation (ICoR) of the stick must indeed be its end on the ground. The angular momentum of the $(i+1)$th stick with respect to its bottom end, immediately after the collision, is $Jl\cos^2\theta$ by the angular impulse-momentum theorem. Therefore, its final angular velocity is

$$\omega = \frac{3J\cos^2\theta}{ml},$$

where we have used the fact that the moment of inertia of a uniform stick of mass m and length l about one of its ends is $I_{end} = \frac{1}{3}ml^2$. By the impulse-momentum theorem, the velocity of the center of mass of the $(i+1)$th stick is

$$v_{CM} = \frac{J\cos\theta - J_f}{m}.$$

In order for the ICoR to be located at its bottom end, $v_{CM} = \frac{\omega l}{2}$. This implies that

$$J_f = J\cos\theta\left(\cos\theta - \frac{3}{2}\right).$$

Note that J_f may be negative as we could have guessed the direction of friction wrongly. Imposing the constraint that $\left|\frac{J_f}{J_N}\right| \leq \mu$, we require

$$\left|1 - \frac{3}{2}\cos\theta\right| \leq \mu$$

for the $(i+1)$th stick to immediately rotate about its bottom end after being collided by the ith stick. For the second part, let the ith stick possess an initial angular velocity ω_0, immediately after it was knocked by the $(i-1)$th stick. Then, its angular velocity immediately before its collision with the $(i+1)$th stick, ω_0', is given by the conservation of energy.

$$\frac{1}{2}I_{end}\omega_0'^2 = \frac{1}{2}I_{end}\omega_0^2 + \frac{1}{2}mgl(1-\cos\theta)$$

$$\omega_0' = \sqrt{\omega_0^2 + \frac{3g(1-\cos\theta)}{l}}.$$

Applying the angular impulse-momentum theorem to the ith stick about its center of mass as it collides with the $(i+1)$th stick, its final angular velocity ω_0'' must obey

$$I_{CM}\omega_0' - \frac{Jl}{2} = I_{CM}\omega_0''$$

$$\omega_0'' = \omega_0' - \frac{6J}{ml},$$

where $I_{CM} = \frac{1}{12}ml^2$ for a uniform rod. It is paramount to understand that there are no impulses delivered by the normal force and friction to the ith stick as the impulse J tends to lift it off the ground — causing the normal force and hence, friction, to be zero. Moving on, notice that the final angular velocity ω of the $(i+1)$th stick after the collision has already been computed previously as $\frac{3J\cos^2\theta}{ml}$. Thus, applying the conservation of energy during the collision,

$$\frac{1}{2}m\left(\frac{l\omega_0'}{2}\right)^2+\frac{1}{2}I_{CM}\omega_0'^2=\frac{1}{2}m\left(\frac{l\omega_0'}{2}-\frac{J}{m}\right)^2+\frac{1}{2}I_{CM}\omega_0''^2+\frac{1}{2}I_{end}\omega^2,$$

where $\frac{l\omega_0'}{2}$ and $\frac{l\omega_0'}{2}-\frac{J}{m}$ are the velocities of the center of the ith stick immediately before and after the collision. Substituting the expressions for ω and ω_0'',

$$J=\frac{2ml\omega_0'}{3\cos^4\theta+4}$$
$$\implies \omega=\frac{3J\cos^2\theta}{ml}=\frac{6\omega_0'\cos^2\theta}{3\cos^4\theta+4}.$$

Substituting the expression for ω_0' in terms of ω_0,

$$\omega=\frac{6\sqrt{\omega_0^2+\frac{3g(1-\cos\theta)}{l}}\cos^2\theta}{3\cos^4\theta+4}.$$

When the subsequent sticks attain the common initial angular velocity ω_f, $\omega=\omega_0=\omega_f$ in the above equation. Solving,

$$\omega_f=\sqrt{\frac{108g\cos^4\theta(1-\cos\theta)}{l(9\cos^8\theta-12\cos^4\theta+16)}}.$$

Actually, ω_f must be the initial angular velocity that we impart to the first stick as well. In principle, we can use ω_f to calculate the time interval between successive collisions and hence the speed of propagation of the "domino wave" (though we shall not do so due to its tedium).

Energy of a System of Interacting Particles

The previous section analyzed the energy of a single particle and rigid body. In this section, we will analyze the energy of a system of interacting particles whose forces of interaction only depend on the relative positions of the two interacting particles. That is, the force on the ith particle due to the jth particle satisfies

$$\boldsymbol{F}_{ij}=\boldsymbol{F}_{ij}(\boldsymbol{r}_i-\boldsymbol{r}_j).$$

Note that the brackets denote "a function of" and not multiplication. We first consider a simple set-up which evokes a rather thought-provoking question.

Problem: Two masses m_1 and m_2 are currently traveling at speeds v_1 and v_2 respectively. If they are connected by a spring with a zero relaxed length that is currently of length x, determine the total conserved mechanical energy of the system of the two masses, while assuming the absence of external forces on both of them.

The spring plays the role of the force of interaction in this case. Now consider the following flawed logic. If we take mass m_1 as a single system and imagine m_2 to be fixed, the total mechanical energy of m_1 should be

$$E_1 = \frac{1}{2}m_1 v_1^2 + \frac{1}{2}kx^2,$$

as this is identical to the case of a particle connected to a spring with a fixed end. Similarly,

$$E_2 = \frac{1}{2}m_2 v_2^2 + \frac{1}{2}kx^2.$$

Then, we erroneously conclude that the total mechanical energy of the system is simply the sum of the two energies:

$$E_{tot} = E_1 + E_2 = \frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 + kx^2.$$

We have in fact double-counted the potential energy — you might expect the reason behind this to be the fact that we cannot imagine the particles to be separately fixed. This is true but the potential energy with a fixed end can in fact be leveraged, as we shall soon see. On a side note, the correct answer is

$$E_{tot} = \frac{1}{2}m_1 v_1^2 + \frac{1}{2}m_2 v_2^2 + \frac{1}{2}kx^2,$$

and we shall understand why in a moment.

Total Energy of a System of Two Particles

To be completely general in the determination of the conserved energy for a system of particles, let us return to the fundamental work-energy theorem for a single particle.

$$W_{ext}^1 + W_{12} = \Delta T_1.$$

That is, the change in kinetic energy of the first particle ΔT_1 is equal to the work W_{ext}^1 done on it by external forces and the work W_{12} done on it by the second particle. Adopting a similar notation,

$$W_{ext}^2 + W_{21} = \Delta T_2.$$

Adding the two equations above,

$$(W_{ext}^1 + W_{ext}^2) + (W_{12} + W_{21}) = \Delta(T_1 + T_2)$$

$$W_{ext} + \int_{\boldsymbol{r}_{10}}^{\boldsymbol{r}_{11}} \boldsymbol{F}_{12} \cdot d\boldsymbol{r}_1 + \int_{\boldsymbol{r}_{21}}^{\boldsymbol{r}_{21}} \boldsymbol{F}_{21} \cdot d\boldsymbol{r}_2 = \Delta T_{tot},$$

where W_{ext} is the total work performed by external forces on the two particles and T_{tot} is the total kinetic energy of the two particles. $\boldsymbol{r}_{10}$ and $\boldsymbol{r}_{20}$ are the initial position vectors of the first and second particles respectively, while $\boldsymbol{r}_{11}$ and $\boldsymbol{r}_{21}$ are their final position vectors. Notice that the two integrals are difficult to integrate separately as $\boldsymbol{F}_{12}$ and $\boldsymbol{F}_{21}$ are dependent on both $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. However, since Newton's third law implies that $\boldsymbol{F}_{21} = -\boldsymbol{F}_{12}$ at every instance in time,

$$W_{ext} + \int_{\boldsymbol{r}_{10}-\boldsymbol{r}_{20}}^{\boldsymbol{r}_{11}-\boldsymbol{r}_{21}} \boldsymbol{F}_{12} \cdot d\left(\boldsymbol{r}_1 - \boldsymbol{r}_2\right) = \Delta T_{tot}.$$

Now, the integral can be evaluated readily if we are given the function $\boldsymbol{F}_{12}$ — remember that $\boldsymbol{F}_{12}$ is a function of $\boldsymbol{r}_1 - \boldsymbol{r}_2$. Suppose that we can associate an internal potential energy U_{12} for this interaction such that

$$U_{12} = -\int_{\boldsymbol{r}_{10}-\boldsymbol{r}_{20}}^{\boldsymbol{r}_{11}-\boldsymbol{r}_{21}} \boldsymbol{F}_{12} \cdot d\left(\boldsymbol{r}_1 - \boldsymbol{r}_2\right).$$

Then,

$$W_{ext} = \Delta\left(T_{tot} + U_{12}\right). \tag{6.13}$$

This is the work-energy theorem for a system of two particles, which states that the work done on the system by external forces is equal to its change in kinetic energy and the internal potential energy associated with the interactions between its constituents. Consider the case where the external forces on the system are conservative, such that a potential energy due to external interactions, U_{ext}, can be defined as

$$U_{ext} = -W_{ext}.$$

Then,

$$T_{tot} + U_{ext} + U_{12} = E \tag{6.14}$$

for some constant E which is the total mechanical energy of the system of two particles. This is the conservation of energy for a system of two particles — the sum of their total kinetic energies, the external potential energy due to external influences and the internal potential energy due to their interactions is conserved if there are no net external non-conservative forces. Finally, let us examine U_{12} in greater detail.

$$U_{12}(\boldsymbol{r}_1 - \boldsymbol{r}_2) = -\int_{\boldsymbol{r}_{10}-\boldsymbol{r}_{20}}^{\boldsymbol{r}_{11}-\boldsymbol{r}_{21}} \boldsymbol{F}_{12}(\boldsymbol{r}_1 - \boldsymbol{r}_2) \cdot d\,(\boldsymbol{r}_1 - \boldsymbol{r}_2)\,.$$

Observe that if we make a substitution of variables $\boldsymbol{r}' = \boldsymbol{r}_1 - \boldsymbol{r}_2$,

$$U_{12}(\boldsymbol{r}') = -\int_{\boldsymbol{r}'_0}^{\boldsymbol{r}'_1} \boldsymbol{F}_{12}(\boldsymbol{r}') \cdot d\boldsymbol{r}'.$$

Clearly, this integral is only dependent on the relative separation of the two particles! Since this integral should be path-independent, we can fix the second particle at the origin and evaluate the potential energy of the first particle at a certain separation $\boldsymbol{r'_1}$ to determine U_{12}. In the set-up concerning two masses attached to a common spring above, U_{12} is thus the potential energy of the first particle due to a spring with a fixed end.

$$U_{12} = \frac{1}{2}kx^2.$$

The total mechanical energy of the previous set-up is then given by Eq. (6.14).

$$E = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 + \frac{1}{2}kx^2.$$

Having understood the loophole in the previous logic, one may ponder if it is valid to apportion the potential energy between the two particles (e.g. $\frac{1}{4}kx^2$ each or $\frac{3}{8}kx^2$ and $\frac{1}{8}kx^2$). The answer is no — U_{12} is the potential energy associated with the interactions between **both** particles and is dependent on both $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, such that even if a particle claims that it owns a certain amount of potential energy, the other particle can always move to receive work and infringe on the previous particle's "possessions". Using the analogy of money again, the joint firm founded by both particles now receives a certain amount of funds from its investors. Whether the money is particle 1's or particle 2's cannot be distinguished as it is a joint venture so it makes no sense to ascribe ownership. Both particles can withdraw money from these funds as and when they want, until it is completely depleted. It just so happens that in the cases where particle 2 is fixed such that its position vector doesn't change, U_{12} will only be dependent on $\boldsymbol{r}_1$ such that we can associate the entire U_{12} with particle 1 — it is simply a fluke for particle 1 as it has found a not-so-greedy partner.

Total Energy of an Arbitrary System of Interacting Particles

Applying the above process to a system of N particles, it is not difficult to prove that the total internal potential energy associated with the interactions between pairs of particles is

$$U_{int} = \sum_{i<j} U_{ij}. \tag{6.15}$$

That is, the internal energy of the system is equal to the sum of the potential energies associated with individual pairs of particles (i, j and j, i are counted as the same pair — the $i < j$ expression under the summation precludes such double-counting). Once again, U_{ij} can be computed by fixing the jth particle at the origin and computing the potential energy associated with the ith particle.

The work-energy theorem for a system of interacting particles states that the external work performed on the system is equal to the change in its total kinetic energy and internal potential energy.

$$W_{ext} = \Delta(T_{tot} + U_{int}). \tag{6.16}$$

The total mechanical energy of a system of particles is then the sum of the total kinetic energy, external potential energy associated with external interactions and the internal potential energy.

$$E = T_{tot} + U_{int} + U_{ext}. \tag{6.17}$$

E is conserved once again if there is no net external non-conservative force on the system. Now, you might be wondering why we could derive the work-energy theorem and the conservation of energy for a rigid body before this section. It just so happens that the forces of interactions between particles on a rigid body often only depend on the magnitudes of their separations — causing the associated potential energies between pairs of particles to follow suit. As the relative distances between particles on a rigid body are preserved, U_{int} is constant in the case of a rigid body, causing the work-energy theorem and the conservation of energy equation above to be reduced to Eqs. (6.8) and (6.10) (where E excludes U_{int}).

Non-conservative Forces

Unfortunately for non-conservative forces, we are unable to define a potential energy function. However, our work-energy theorem is still valid and enables us to derive certain useful results. The total work done by external forces can be divided into two components: the work done by conservative forces,

W_{con}, and by non-conservative forces, W_{noncon}. For a single particle or rigid body,

$$W_{con} + W_{noncon} = \Delta T.$$

Since $W_{con} = -\Delta U_{ext}$,

$$W_{noncon} = \Delta(T + U_{ext}).$$

For a system of particles,

$$\begin{aligned} W_{con} + W_{noncon} &= \Delta(T_{tot} + U_{int}) \\ W_{noncon} &= \Delta(T + U_{int} + U_{ext}). \end{aligned}$$

In all cases,

$$W_{noncon} = \Delta E.$$

The work done by non-conservative forces is equal to the change in the mechanical energy of the system they act on.

6.4 Deriving the Equation of Motion from Energy

For a system whose energy is only dependent on a single coordinate, its equation of motion can be obtained by differentiating the conservation of energy equation. Consider the following problem.

Problem: A uniform sphere, of mass m and radius R, rolls down an inclined plane with an angle of inclination, θ. Find the sphere's acceleration, assuming that it rolls without slipping and that the plane remains stationary.

Instead of analyzing forces and torques like we did in the previous chapter, we can employ the fact that the total mechanical energy of the sphere is conserved as static friction does no work (there is no relative motion at the point of contact by the non-slip condition). The total mechanical energy E

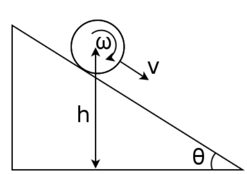


Figure 6.5: Sphere on inclined plane

at one instant in time is

$$E = \frac{1}{2}mv^2 + \frac{1}{2}I\omega^2 + mgh,$$

where the instantaneous velocity of the center of the sphere, v, is in the direction parallel to the plane (positive downwards) and h is the vertical position of the center, as measured with respect to the bottom of the plane. The instantaneous angular velocity ω is defined to be positive in the clockwise direction. There are then two ways to proceed from here.

Method 1: Since the total mechanical energy is conserved, its rate of change is zero, $\frac{dE}{dt} = 0$.

$$mav + \frac{2}{5}mr^2\alpha\omega + mg\frac{dh}{dt} = 0,$$

where a is also in the direction parallel to the plane and α is the angular acceleration clockwise. Next we observe that

$$\frac{dh}{dt} = -v\sin\theta,$$

as $\frac{dh}{dt}$ is the vertical velocity of the center of the sphere (positive in the upward direction) while v is the velocity of the sphere in the downward direction parallel to the plane. Lastly, using the non-slip condition $v = r\omega$ and $a = r\alpha$, we can write

$$mav + \frac{2}{5}mav - mgv\sin\theta = 0.$$

Cancelling the v's,

$$a = \frac{5}{7}g\sin\theta.$$

Note that we can cancel the v's regardless of whether its magnitude is zero for the following physical reason: the dynamical laws do not depend on the current velocity of the sphere. Thus, regardless of whether the sphere just begins moving or already possesses a certain velocity, it will experience the same acceleration.

Method 2: Let s be the distance of the point of contact of the sphere with the plane, from the top of the plane. Then, $v = \dot{s}$ and we can write

$$E = \frac{1}{2}m\dot{s}^2 + \frac{1}{5}m\dot{s}^2 + mgh.$$

Then, by taking $\frac{dE}{ds} = 0$,

$$\frac{7}{10}m\frac{d\dot{s}^2}{ds} + mg\frac{dh}{ds} = 0.$$

Observing that $\frac{d\dot{s}^2}{ds} = 2\ddot{s}$ and $\frac{dh}{ds} = -\sin\theta$,

$$\frac{7}{5}m\ddot{s} = mg\sin\theta$$
$$\ddot{s} = \frac{5}{7}g\sin\theta.$$

The two methods are essentially the same, except that the second method adroitly circumvents the questionable boundary case when $v = 0$.

6.5 Galilean Transformations and Center of Mass Frame

Before we apply the conservation laws in the previous sections in solving new types of problems, let us digress for a while and discuss about Galilean transformations and the center of mass frame which will be useful later.

6.5.1 *Galilean Transformations*

The Newtonian formulation assumes the existence of absolute space and time. Inertial frames move at a constant velocity with respect to this absolute space and share the same universal time. A Galilean transformation is a transformation from one inertial frame to another under these assumptions of space-time. Formally, consider two inertial frames S and S'. An event in frame S occurs at position (x, y, z) and time t. Now suppose that frame S' is moving at a velocity $\boldsymbol{v}$ with respect to frame S (Fig. 6.6). We wish to derive the coordinates (x', y', z') and t' of the same event as measured by an observer in S'. We shall append a prime, $'$, behind our quantities to denote that they are measured with respect to S'.

Based on the second axiom regarding a universal time,

$$t' = t.$$

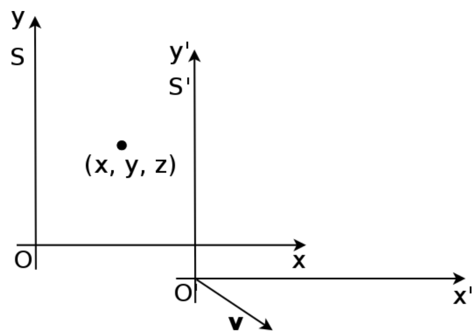


Figure 6.6: Galilean transformation

Furthermore, the discrepancy in the coordinates of the event as measured by S' originates from its own motion. Assuming that the origins of the coordinate frames coincided at $t = 0$,

$$\begin{aligned} x' &= x - v_x t, \\ y' &= y - v_y t, \\ z' &= z - v_z t, \\ \implies \boldsymbol{r}' &= \boldsymbol{r} - \boldsymbol{v}t, \end{aligned}$$

where $\boldsymbol{r}$ and $\boldsymbol{r}'$ are the position vectors of the event as measured by an observer in S and S' respectively. Let the velocity of a particle with respect to frame S be $\boldsymbol{u}$. Then the velocity of that particle with respect to S', $\boldsymbol{u}'$, is

$$\boldsymbol{u}' = \frac{d\boldsymbol{r}'}{dt} = \frac{d\boldsymbol{r}}{dt} - \boldsymbol{v} = \boldsymbol{u} - \boldsymbol{v},$$

which is a very intuitive result which states that we can simply subtract the relative velocity of a frame with respect to another to calculate the velocity of a particle as observed in the former frame. Moreover, since the mass of a particle is assumed to be invariant across inertial frames (as it is deemed an intrinsic property), we can then relate its momentum with respect to S', denoted by $\boldsymbol{p'}$, to that with respect to S, which we will call $\boldsymbol{p}$.

$$\begin{aligned} \boldsymbol{p'} &= m(\boldsymbol{u} - \boldsymbol{v}) = \boldsymbol{p} - m\boldsymbol{v} \\ \implies \Delta\boldsymbol{p'} &= \Delta\boldsymbol{p}. \end{aligned}$$

Summing these individual values for a system of particles, it implies that the change in momentum of a system as measured by S is the same as that measured by S'. Thus, if the total momentum of a system is conserved in S, it is also conserved in S'.

Now let us consider the acceleration of a particle in S, indicated by $\boldsymbol{a}$, and in S', represented by $\boldsymbol{a'}$.

$$\boldsymbol{a'} = \frac{d\boldsymbol{u'}}{dt} = \frac{d\boldsymbol{u}}{dt} = \boldsymbol{a}.$$

This means that the acceleration of a particle is the same in both frames. A last assumption of Galilean relativity is that forces are invariant. Therefore, if the net forces in S and S' are $\boldsymbol{F}$ and $\boldsymbol{F}'$ respectively,

$$\boldsymbol{F} = \boldsymbol{F}'.$$

Since

$$\begin{aligned} \boldsymbol{F} &= m\boldsymbol{a} = m\boldsymbol{a}' \\ \implies \boldsymbol{F}' &= m\boldsymbol{a}'. \end{aligned}$$

That is, Newton's second law is valid in all inertial frames. Lastly, we wish to derive an expression for the kinetic energy of a system, T', in S' but we shall first introduce the center of mass frame as it has certain unique properties.

6.5.2 *Center of Mass*

As stated before, the position vector $\boldsymbol{R}$ of the center of mass of a system of N discrete particles or a continuous mass distribution is defined as

$$\boldsymbol{R} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{r}_i}{\sum_{i=1}^{N} m_i} = \frac{\int \boldsymbol{r} dm}{\int dm}.$$

The velocity of the center of mass, $\boldsymbol{v}_{CM}$ is then

$$\boldsymbol{v}_{CM} = \frac{\sum_{i=1}^{N} m_i \boldsymbol{v}_i}{\sum_{i=1}^{N} m_i} = \frac{\int \boldsymbol{v} dm}{\int dm}.$$

Note that this implies that the center of mass of a system travels at a constant velocity $\boldsymbol{v}_{CM}$ if there are no net external forces as the total linear momentum is conserved.

The Center of Mass Frame

Following from the expressions above, let us consider the motion of a system with respect to a frame attached to the center of mass of the system. For a discrete system (the following can be easily extended to the continuous case), let the velocity of the ith particle in the center of mass frame be $\boldsymbol{u}_i'$ and that in the lab frame be $\boldsymbol{u}_i$. Then,

$$\begin{aligned}\boldsymbol{u}_i' &= \boldsymbol{u}_i - \boldsymbol{v}_{CM},\\ \boldsymbol{p'} &= \sum_{i=1}^{N} m_i \boldsymbol{u}_i' \\ &= \sum_{i=1}^{N} m_i(\boldsymbol{u}_i - \boldsymbol{v}_{CM}) \\ &= M\boldsymbol{v}_{CM} - M\boldsymbol{v}_{CM},\end{aligned}$$

where M is the total mass of the system. Then,

$$\boldsymbol{p}' = 0. \tag{6.18}$$

Thus, the total momentum of the system with respect to the center of mass frame, $\boldsymbol{p}'$, is 0.

6.5.3 *Kinetic Energy Transformation*

We now wish to relate the kinetic energy T' of a system of particles with respect to frame S', to T, its kinetic energy with respect to frame S. Once again, frame S' is traveling at velocity $\boldsymbol{v}$ with respect to frame S. Adopting the same notation as before,

$$\begin{aligned}T' &= \frac{1}{2}\sum_{i=1}^{N} m_i \boldsymbol{u}_i' \cdot \boldsymbol{u}_i' \\ &= \frac{1}{2}\sum_{i=1}^{N} m_i (\boldsymbol{u}_i - \boldsymbol{v}) \cdot (\boldsymbol{u}_i - \boldsymbol{v}) \\ &= \frac{1}{2}\sum_{i=1}^{N} m_i u_i^2 + \frac{1}{2}\sum_{i=1}^{N} m_i v^2 - \sum_{i=1}^{N} m_i \boldsymbol{u}_i \cdot \boldsymbol{v} \\ &= T + \frac{1}{2}Mv^2 - \boldsymbol{v} \cdot \sum_{i=1}^{N} m_i \boldsymbol{u}_i.\end{aligned}$$

This expression is not particularly edifying in itself. However, if we choose frame S to be our center of mass frame, $\sum_{i=1}^{N} m_i \boldsymbol{u}_i = 0$. Thus,

$$T' = T_{CM} + \frac{1}{2}Mv_{CM}^2, \tag{6.19}$$

where v_{CM} is the speed of the center of mass in frame S'. We see that the kinetic energy of a system viewed in a frame S' is simply the sum of the kinetic energy of that system in the center of mass frame and the kinetic energy by treating the system as a single mass M traveling at v_{CM}. Furthermore, by taking the changes of both sides, it can be seen that the **changes** in the kinetic energies of a system with respect to all inertial frames are identical if there is no net external force acting on it (so that v_{CM} is constant). Moreover, since the potential energy of a system of particles only depends on their relative positions, the potential energy of a system is invariant across inertial frames. Ultimately, this means that if there are no net external forces and the total mechanical energy of a system is conserved in one inertial frame, the total mechanical energy is conserved in all inertial frames.

6.6 Collisions

The collisions between objects constitute a typical class of problems. Generally, there can be two classifications — elastic and inelastic collisions. The former refers to the situation where the total kinetic energy of the system of colliding bodies is conserved while the latter means that some of that energy is lost during the collision. Collision problems can generally be solved easily using the conservations of linear and angular momenta, and predetermined conditions on the energy of the system (such as the condition that 40% of the system's kinetic energy is lost after collision).

6.6.1 *Elastic Collisions*

One-dimensional Collisions

In a one-dimensional elastic collision, two particles initially traveling at certain velocities along a line collide and separate with certain velocities along the same line. In such cases, we can apply the following theorem instead of the conservation of energy principle, which involves a cumbersome quadratic equation.

Theorem: The final relative velocity between two particles is the negative of the initial relative velocity in a 1-D elastic collision.

Proof: We define u_i to be the initial velocity of the ith particle and v_i to be the final velocity of the ith particle. We will adopt this definition throughout this chapter. Applying the conservation of momentum and energy to the system comprising both particles, in light of the lack of external forces,

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2,$$
$$\frac{1}{2}m_1u_1^2 + \frac{1}{2}m_2u_2^2 = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2.$$

Rearranging,

$$m_1(u_1 - v_1) = m_2(v_2 - u_2),$$
$$\frac{1}{2}m_1(u_1 - v_1)(u_1 + v_1) = \frac{1}{2}m_2(v_2 - u_2)(v_2 + u_2).$$

Dividing the second equation by the first and simplifying, we obtain

$$u_1 - u_2 = v_2 - v_1,$$

which is the statement we wish to prove. We can use this fact, in replacement of the conservation of energy, in combination with the conservation of linear momentum to solve such one-dimensional elastic collisions between two particles.

Elastic Collisions with Massive Object

Consider the head-on elastic collision of a gas particle, with a small mass m and initial velocity u_1, and the stationary wall of a gas piston with mass M where $m \ll M$. Find the final velocity of the particle v_1.

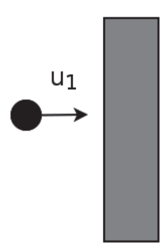


Figure 6.7: Collision of gas particle with massive wall

Using both the equation we just derived and the conservation of momentum,

$$u_1 = v_2 - v_1,$$
$$mu_1 = mv_1 + Mv_2$$

where v_1 and v_2 are the final velocities of the particle and wall, positive in the direction of u_1. Solving, we get

$$v_1 = \frac{m - M}{m + M} u_1,$$
$$v_2 = \frac{2m}{m + M} u_1.$$

As $\frac{m}{M} \to 0$,

$$v_1 = -u_1,$$
$$v_2 = 0.$$

We see that the particle just reverses its direction and travels at the same initial speed while the wall remains practically stationary. Now let us consider the case where the wall initially moves at a velocity u_2 in a direction parallel to u_1. Instead of rewriting the equations again, let us consider a frame that is traveling at a constant velocity u_2. In that frame, the wall is stationary while the particle is initially traveling at speed $u_1 - u_2$. Thus, we are back to the above problem since the energy and momentum of this system are still conserved after a Galilean transformation. Applying the result we derived before, we see that the final velocity of the particle with respect to this frame is simply $u_2 - u_1$. Thus, the final velocity of the particle in the lab frame, which is the sum of the final velocity as viewed by the moving frame and the velocity of the moving frame relative to the lab frame, is $2u_2 - u_1$. Interestingly, the change in the kinetic energy of the particle is

$$\Delta T = T_f - T_i = \frac{1}{2} m (2u_2 - u_1)^2 - \frac{1}{2} m u_1^2 = 2(u_2^2 - u_1 u_2).$$

Observe that when u_1 and u_2 are opposite in direction, the total kinetic energy of the particle increases. When u_1 and u_2 are in the same direction and $u_1 > u_2$, which is valid most of the time for a gas piston, the kinetic energy of the particle actually decreases. This is the microscopic reason behind the gain and loss in the internal energy of a gas when a piston is pushed or pulled to compress or expand a gas, respectively.

General One-Dimensional Collision

Problem: Two particles, with masses m_1 and m_2, travel along a line with velocities u_1 and u_2. They then collide and still travel along the same line after the collision. Find their final velocities v_1 and v_2.

Considering the conservation of momentum and the relative velocities of the particles,

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2,$$
$$u_1 - u_2 = v_2 - v_1.$$

Solving,

$$v_1 = \frac{2m_2u_2 + (m_1 - m_2)u_1}{m_1 + m_2},$$
$$v_2 = \frac{2m_1u_1 + (m_2 - m_1)u_2}{m_1 + m_2}.$$

A special case occurs when $m_1 = m_2$ — the particles simply exchange their velocities.

Two-Dimensional Elastic Collisions

Consider the following collision (Fig. 6.8). A particle of mass m and initial velocity u_1 undergo a side-on collision with another stationary particle of mass m. Their final velocities are not along the same line. Given θ, find ϕ.

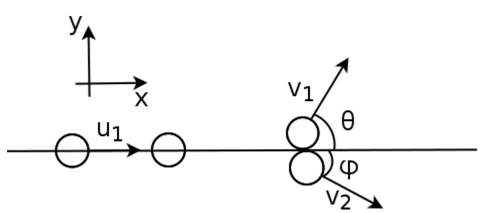


Figure 6.8: 2-D elastic collision

Define the xy-plane to be the plane the particles lie in after their collision. Writing our equations based on the conservation of energy and momentum in the x and y directions,

$$u_1^2 = v_1^2 + v_2^2,$$
$$u_1 = (v_1 \cos\theta + v_2 \cos\phi) \implies u_1^2 = v_1^2 \cos^2\theta + v_2^2 \cos^2\phi + 2v_1v_2 \cos\theta \cos\phi,$$
$$0 = v_1 \sin\theta - v_2 \sin\phi \implies 0 = v_1^2 \sin^2\theta + v_2^2 \sin^2\phi - 2v_1v_2 \sin\theta \sin\phi.$$

Adding the last two equations,

$$u_1^2 = v_1^2 + v_2^2 + 2v_1v_2 \cos(\theta + \phi).$$

Subtracting the above by the first equation,

$$\cos(\theta + \phi) = 0 \implies \phi = \frac{\pi}{2} - \theta.$$

Thus, we see that the particles travel at right angles to each other. We could have also proven this result in a much less tedious way using vectors. By the conservation of momentum and energy,

$$\boldsymbol{u}_1 = \boldsymbol{v}_1 + \boldsymbol{v}_2,$$
$$\boldsymbol{u}_1 \cdot \boldsymbol{u}_1 = \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 + \boldsymbol{v}_2 \cdot \boldsymbol{v}_2.$$

Taking the dot product of the first equation with itself and subtracting the second, we obtain the result

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = 0,$$

which implies that the particles travel at right angles with respect to each other after the collision. We now understand the physics behind billiards!

The Center of Mass Frame

Elastic collisions between two particles are often easy to deal with in the center of mass frame. We shall first prove a property with regard to a collision in the center of mass frame. Let the initial momenta of the two objects with masses m_1 and m_2 be $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ in the center of mass frame, respectively. Let their final momenta be $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$. Then by definition of the center of mass frame,

$$\boldsymbol{p}_1 + \boldsymbol{p}_2 = 0 \implies \boldsymbol{p}_1 = -\boldsymbol{p}_2,$$
$$\boldsymbol{p}_3 + \boldsymbol{p}_4 = 0 \implies \boldsymbol{p}_3 = -\boldsymbol{p}_4.$$

Furthermore, by the conservation of energy,

$$\frac{\boldsymbol{p}_1 \cdot \boldsymbol{p}_1}{2m_1} + \frac{\boldsymbol{p}_2 \cdot \boldsymbol{p}_2}{2m_2} = \frac{\boldsymbol{p}_3 \cdot \boldsymbol{p}_3}{2m_1} + \frac{\boldsymbol{p}_4 \cdot \boldsymbol{p}_4}{2m_2}.$$

Since $\boldsymbol{p}_2 = -\boldsymbol{p}_1$ and $\boldsymbol{p}_4 = -\boldsymbol{p}_3$,

$$\frac{\boldsymbol{p}_1 \cdot \boldsymbol{p}_1}{2m_1} + \frac{-\boldsymbol{p}_1 \cdot -\boldsymbol{p}_1}{2m_2} = \frac{\boldsymbol{p}_3 \cdot \boldsymbol{p}_3}{2m_1} + \frac{-\boldsymbol{p}_3 \cdot -\boldsymbol{p}_3}{2m_2}.$$

Thus,

$$p_1 = p_3,$$
$$p_2 = p_4.$$

We see that for the two particles, the magnitudes of their momenta, and thus the magnitudes of their velocities, do not change after the collision, though their directions may vary.

General Two-Dimensional Collisions

Let us use the center of mass frame to analyze a general two-dimensional collision. A particle m_1 approaches a stationary particle m_2 at velocity u in the positive x-direction and undergoes an off-center collision. Our objective is to analyze the resultant motion of these particles. Note that this problem is rather general as we can simply switch to the frame of m_2 and apply the results of this problem in cases where m_2 moves in the lab frame.

As we shall see, there is, in fact, still one variable parameter as the system is indeterminate based on the above conditions. Consider the center of mass frame which travels at

$$v_{CM} = \frac{m_1 u}{m_1 + m_2}$$

in the x-direction, with respect to the lab frame. The velocities of the two particles in this center of mass frame (Fig. 6.9) are

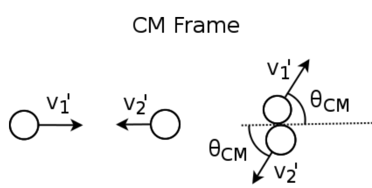


Figure 6.9: Motion in CM frame

$$v_1' = u - v_{CM} = \frac{m_2 u}{m_1 + m_2},$$
$$v_2' = -\frac{m_1 u}{m_1 + m_2},$$

in the x-direction (the x-axis of the center of mass frame is aligned with that of the lab frame). After the collision, the speeds of the particles — $|v_1'| = \frac{m_2 u}{m_1 + m_2}$ and $|v_2'| = \frac{m_1 u}{m_1 + m_2}$ — are preserved but may be directed at arbitrary angles relative to the x-axis. Let the final velocity of m_1 be directed at θ_{CM} above the positive x-axis, in the center of mass frame. This shall be the only parameter in this system.

Let the final vectorial velocities of the particles in the center of mass frame be $\boldsymbol{v}_1'$ and $\boldsymbol{v}_2'$, respectively. Then, the final velocity of each particle in the lab frame is the sum of its final vectorial velocity, as measured in the center of mass frame, and the velocity of the center of mass in the lab frame

$\boldsymbol{v}_{CM}$. That is, for m_1,

$$\boldsymbol{v}_1 = \boldsymbol{v}_1' + \boldsymbol{v}_{CM} = \frac{u}{m_1 + m_2} \begin{pmatrix} m_1 + m_2 \cos\theta_{CM} \\ m_2 \sin\theta_{CM} \end{pmatrix}.$$

For the total momentum to be nullified in the center of mass frame, the velocity of m_2 in the center of mass frame must be the negative of the velocity of m_1 in the center of mass frame, scaled by a factor of $\frac{m_1}{m_2}$.

$$\boldsymbol{v}_2' = -\frac{m_1}{m_2}\boldsymbol{v}_1'.$$

Thus,

$$\boldsymbol{v}_2 = \boldsymbol{v}_2' + \boldsymbol{v}_{CM} = \frac{u}{m_1 + m_2} \begin{pmatrix} m_1 - m_1 \cos\theta_{CM} \\ -m_1 \sin\theta_{CM} \end{pmatrix}.$$

There are several interesting results from the above analysis. Firstly, consider the possible angles of deflection in the lab frame for m_1.

We have to consider two cases — namely, when $v_{CM} \geq v_1'$ and $v_{CM} < v_1'$. The former case occurs when $m_1 \geq m_2$ while the latter occurs when $m_1 < m_2$. For both cases, let $\overline{OO'}$ denote the vector $\boldsymbol{v}_{CM}$ and draw a circle of radius v_1' about O'. $\boldsymbol{v}_1$ is the sum of $\boldsymbol{v}_{CM}$ and $\boldsymbol{v}_1'$. The latter vector can be directed from O' to any point on the circle (as θ_{CM} is not fixed).

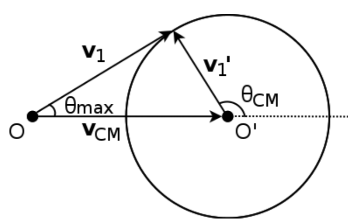


Figure 6.10: Case 1: $v_{CM} \geq v_1'$

In the first scenario, O lies outside or on the boundary of the circle — imposing an upper bound on the possible angle of deflection. From Fig. 6.10 above, it is evident that the maximum angle of deflection in the lab frame, θ_{max} occurs when $\boldsymbol{v}_1$ and $\boldsymbol{v}_1'$ are perpendicular.

$$\theta_{max} = \sin^{-1}\frac{v_1'}{v_{CM}} = \sin^{-1}\frac{m_2}{m_1}.$$

The final speed of m_1 is then

$$v_1 = \frac{\sqrt{m_1^2 - m_2^2}}{m_1 + m_2}u.$$

The final velocities are

$$\boldsymbol{v}_1 = \begin{pmatrix} \frac{m_1 - m_2}{m_1} u \\ \frac{\sqrt{m_1^2 - m_2^2} m_2}{(m_1 + m_2) m_1} u \end{pmatrix},$$

$$\boldsymbol{v}_2 = \begin{pmatrix} u \\ -\frac{\sqrt{m_1^2 - m_2^2}}{m_1 + m_2} u \end{pmatrix}.$$

Furthermore, observe that for angles of deflection below θ_{max}, there are two possible final configurations and hence, two possible final velocities that correspond to one angle of deflection (a line drawn with a smaller angle of deflection has two intersections with the circle).

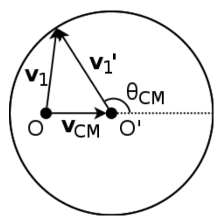


Figure 6.11: Case 2: $v_{CM} < v_1'$

When $v_{CM} < v_1'$ (i.e. $m_1 < m_2$), the point O lies within the circle (Fig. 6.11). Then, all angles of deflection are possible. By considering the two cases above, it can be seen that the backscattering of m_1 (such that the horizontal component of its final velocity is in the negative x-direction) is only possible if $m_1 < m_2$. For particle m_2, a similar analysis follows, with the provision that $v_{CM} = |v_2'|$. Thus, O always lies on the boundary of the circle around O' for m_2 — implying that particle m_2 can never be backscattered.

Angle Between Resultant Velocities

The angle θ_{sep} between the resultant velocities of the particles in the lab frame can be computed via the dot product of $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$.

$$\begin{aligned} \cos\theta_{sep} &= \frac{\boldsymbol{v}_1 \cdot \boldsymbol{v}_2}{|\boldsymbol{v}_1||\boldsymbol{v}_2|} \\ &= \frac{m_1^2 u^2 - m_1 m_2 u^2 + (m_1 m_2 - m_1^2) u^2 \cos\theta_{CM}}{\sqrt{m_1^2 u^2 + 2 m_1 m_2 u^2 \cos\theta_{CM} + m_2^2 u^2} \cdot \sqrt{m_1^2 u^2 - 2 m_1^2 u^2 \cos\theta_{CM} + m_1^2 u^2}} \\ &= \frac{(m_1 - m_2)\sqrt{1 - \cos\theta_{CM}}}{\sqrt{2(m_1^2 + m_2^2 + 2 m_1 m_2 \cos\theta_{CM})}}. \end{aligned}$$

It can be seen that $\theta_{sep} = \frac{\pi}{2}$, when $m_1 = m_2$, is merely a special case of the above result.

6.6.2 *Inelastic Collisions*

Perfectly Inelastic Collisions

A perfectly inelastic collision is a collision where the greatest loss of kinetic energy is observed. We shall now prove that this happens when the particles "stick" together after collision (i.e. they travel at the same final velocity).

Adopting the same notation for the initial and final velocities of the particles, we attempt to minimize the final kinetic energy of the system, which is

$$T_f = \sum_{i=1}^{N} \frac{1}{2} m_i v_i^2.$$

This can be related to the final kinetic energy in the center of mass frame

$$T_f = T_{CM} + \frac{1}{2} M v_{CM}^2,$$

as proven earlier. It is obvious that the kinetic energy of a system, with respect to any frame of reference, is always greater than or equal to zero. Since v_{CM} is constant as a consequence of the conservation of linear momentum, T_f is minimum when T_{CM} is 0. In order for this to be true, all particles must be stationary in the center of mass frame after the collision — meaning that they travel at the same resultant velocity v_{CM} in the lab frame.

Problem: A particle of mass m_1 of initial velocity u_1 collides and sticks to a particle of mass m_2 and initial velocity u_2, which is along the same direction as u_1. Determine the energy lost in this inelastic collision.

We can again consider the center of mass frame, in which the initial velocities of the particles are

$$u_1' = \frac{m_2(u_1 - u_2)}{m_1 + m_2},$$

$$u_2' = \frac{m_1(u_2 - u_1)}{m_1 + m_2}.$$

Remember that the change in kinetic energy is invariant across all inertial frames if there is no net external force on the system (such that v_{CM} is constant). Therefore, we can determine the total loss in kinetic energy in the center of mass frame, which is the total initial energy. This is given by

$$\frac{1}{2}m_1u_1'^2 + \frac{1}{2}m_2u_2'^2 = \frac{1}{2}\frac{m_1m_2}{m_1+m_2}(u_1-u_2)^2,$$

as the final velocities of the particles are zero in the center of mass frame.

Inelastic Collisions

In the more general case, an inelastic collision may occur such that some energy is lost, but it may be less than or equal to that in the perfectly inelastic case. Then, the coefficient of restitution, e, of a two-particle collision supersedes the conservation of energy equation. e is defined to be the magnitude of the ratio of the final relative speed of the two particles to their initial relative speed, usually along the line of impact.

$$e = \frac{v_{2x}-v_{1x}}{u_{1x}-u_{2x}},$$

where the x-direction has been defined to be the line of impact. For a one-dimensional collision,

$$e = \frac{v_2-v_1}{u_1-u_2},$$

$e = 1$ corresponds to an elastic collision while $e = 0$ corresponds to a perfectly inelastic collision. We can analyze a collision associated with a coefficient of restitution e in the center of mass frame, while adopting the same definitions for $\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$. Due to the definition of the center of mass frame,

$$\boldsymbol{p}_2 = -\boldsymbol{p}_1,$$
$$\boldsymbol{p}_4 = -\boldsymbol{p}_3.$$

From the definition of the coefficient of restitution, and because relative velocities do not vary across inertial frames,

$$e = \frac{v_{2x}-v_{1x}}{u_{1x}-u_{2x}} = \frac{\frac{p_{4x}}{m_2}-\frac{p_{3x}}{m_1}}{\frac{p_{1x}}{m_1}-\frac{p_{2x}}{m_2}} = \frac{-\left(\frac{1}{m_1}+\frac{1}{m_2}\right)p_{3x}}{\left(\frac{1}{m_1}+\frac{1}{m_2}\right)p_{1x}}$$

$$\implies p_{3x} = -ep_{1x}.$$

Similarly,

$$p_{4x} = -ep_{2x}.$$

The above results imply that the component of the final velocities along the line of impact is e times that of the negated initial velocities in the center of mass frame. For a one-dimensional collision,

$$v_1' = -eu_1',$$
$$v_2' = -eu_2',$$

where a prime has been used to denote that these quantities are measured in the center of mass frame.

Problem: A particle of mass m_1 and initial velocity u_1 undergoes a head-on collision with another particle of mass m_2 and initial velocity u_2. If the coefficient of restitution is e, determine the final velocities of the particles and the total loss in energy due to the collision.

In the center of mass frame, which travels at velocity $v_{CM} = \frac{m_1u_1+m_2u_2}{m_1+m_2}$ with respect to the lab frame, the initial velocities of the particles are

$$u_1' = \frac{m_2(u_1 - u_2)}{m_1 + m_2},$$
$$u_2' = \frac{m_1(u_2 - u_1)}{m_1 + m_2}.$$

The final velocities of the particles in the center of mass frame are

$$v_1' = -eu_1' = \frac{em_2(u_2 - u_1)}{m_1 + m_2},$$
$$v_2' = -eu_2' = \frac{em_1(u_1 - u_2)}{m_1 + m_2}$$

Thus, the final velocities of the particles in the lab frame are

$$v_1 = v_{CM} + v_1' = \frac{em_2(u_2 - u_1) + m_1u_1 + m_2u_2}{m_1 + m_2},$$
$$v_2 = v_{CM} + v_2' = \frac{em_1(u_1 - u_2) + m_1u_1 + m_2u_2}{m_1 + m_2}.$$

The total loss in energy is the same in all inertial frames as the total momentum of the system is conserved. Thus, we can simply compute the loss in the center of mass frame which is

$$\frac{1}{2}m_1u_1'^2 + \frac{1}{2}m_2u_2'^2 - \frac{1}{2}m_1v_1'^2 - \frac{1}{2}m_2v_2'^2 = \frac{1}{2}(1-e^2)m_1u_1'^2 + \frac{1}{2}(1-e^2)m_2u_2'^2$$
$$= \frac{(1-e^2)m_1m_2(u_1 - u_2)^2}{2(m_1 + m_2)}.$$

6.6.3 *Collisions with a Rigid Body*

In a collision between a particle and a rigid body, the principle of the conservation of angular momentum, with respect to an arbitrary fixed point, should be leveraged in addition to that of momentum conservation. The point of collision or a point lying on the line of impact generally functions as a convenient origin.

Problem: A point bead of mass m travels at speed u and collides elastically with a stationary uniform rod, of mass $3m$ and length l, at a point that is $\frac{l}{4}$ distance away from its center of mass. The center of masses of both objects travel in the direction of the the initial velocity of the bead after the collision. What are the final velocities of the bead and the center of mass of the rod, and the angular velocity of the rod?

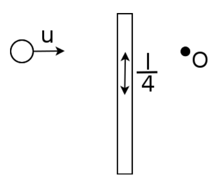


Figure 6.12: Colliding bead and rod

Analyzing this problem, we see that angular momentum, linear momentum and the total energy of this system are conserved as there is no net external force on this system. Let the final velocities of the bead and the center of mass of the rod be v_1 and v_2 respectively, and let the final clockwise angular velocity of the rod be ω. We pick a stationary origin, O, that is at the same vertical level as the bead for our angular momentum calculations so that the angular momentum due to the bead is zero, both before and after the collision.

$$mu = mv_1 + 3mv_2, \tag{CoLM}$$

$$0 = I_{CM}\omega + M\left(\boldsymbol{r}_{CM} \times \boldsymbol{v}_{CM}\right)_z = \frac{1}{4}ml^2\omega - \frac{3}{4}mlv_2, \tag{CoAM}$$

$$\frac{1}{2}mu^2 = \frac{1}{2}mv_1^2 + \frac{3}{2}mv_2^2 + \frac{1}{8}ml^2\omega^2, \tag{CoE}$$

where $I_{CM} = \frac{1}{12} \cdot 3m \cdot l^2 = \frac{1}{4}ml^2$. From the second equation,

$$3v_2 = l\omega.$$

Substituting this into the first and third equations,

$$mu = mv_1 + ml\omega,$$
$$mu^2 = mv_1^2 + \frac{7}{12}ml^2\omega^2.$$

We can isolate and eliminate v_1, so that

$$(u - l\omega)^2 = u^2 - \frac{7}{12}l^2\omega^2$$
$$\omega = \frac{24u}{19l},$$
$$v_2 = \frac{8}{19}u,$$
$$v_1 = -\frac{5}{19}u.$$

6.7 Varying Amounts of Moving Mass

You and your $(N-1)$ friends have just come up with a brilliant idea — a human-propelled boat! Coincidentally, you all possess the same mass m. Initially all N of you stand on a stationary boat which has a mass assumed to be negligible. Due to the limited strength of you and your friends, each of you can only run on the boat and jump off at a velocity u with respect to the boat. How should $(N-1)$ people jump to propel the last person at the greatest velocity? The boat rests on frictionless ground.

Suppose that all $(N-1)$ people jump off at the boat at the same time. Let the final velocity of the boat be v. Then, by the conservation of momentum,

$$0 = (N-1)m(v-u) + mv \implies v = \frac{N-1}{N}u.$$

However, now suppose that the $(N-1)$ people jump off one at a time. Let v_i be the speed of the boat after i people have jumped off. Consider the event where the $(i+1)$th person jumps off. Before the jump, the total momentum of the boat and the passengers on the boat is $(N-i)mv_i$. After the jump, the boat travels at v_{i+1} — implying that the person who leapt off the boat travels at $v_{i+1} - u$ in the lab frame (Fig. 6.13).

By the conservation of momentum,

$$(N-i)mv_i = (N-i-1)mv_{i+1} + m(v_{i+1} - u).$$

Thus,

$$v_{i+1} = v_i + \frac{1}{N-i}u.$$

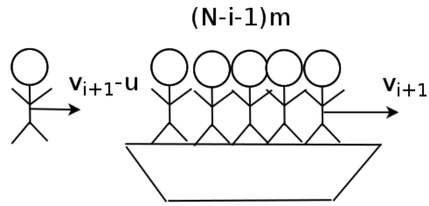


Figure 6.13: Propelling a boat

Repeatedly applying this recursion formula with $v_0 = 0$,

$$v_{N-1} = \left(\frac{1}{2} + \cdots + \frac{1}{N}\right) u = u \sum_{i=2}^{N} \frac{1}{i}.$$

We observe that this is larger than the previous final velocity as there are $(N-1)$ terms which are each individually equal to or larger than $\frac{u}{N}$. The second solution is, in fact, the optimal solution. This is because the velocity of the people that have jumped off the boat in the lab frame is $-\boldsymbol{u}$ relative to the resultant velocity of the boat after the jump. If they had jumped off in intervals, each individual would have left the boat at a larger velocity towards the left in the lab frame as compared to the situation where they simultaneously jumped, as the boat has not sped up yet. Since the total momentum of the system of people must still be conserved, the final person and boat will be propelled at a larger velocity.

Finally, suppose that there are now $j(N-1)$ people with mass $\frac{m}{j}$ while the final person has mass m (so that the total mass remains the same). Repeating the above calculations where the people jump off one at a time, the final velocity of the last person on the boat is

$$\begin{aligned} v_{jN-1} &= \left(\frac{1}{j+1} + \cdots + \frac{1}{jN}\right) u \\ &= u \sum_{i=j+1}^{jN} \frac{1}{i} > \left(j \cdot \frac{1}{2j} + j \cdot \frac{1}{3j} + \cdots + j \cdot \frac{1}{jN}\right) u, \end{aligned}$$

where the inequality is obtained from dividing consecutive terms into groups of length j and where the last expression is v_{N-1} derived previously. It is, therefore, also beneficial to divide the propulsion of masses into smaller bits, in addition to thrusting them individually.

Fuel-Propelled Rocket

Encouraged by the fact that interspersed propulsions of small masses optimize the speed of the boat, let us consider the case where we replace the boat

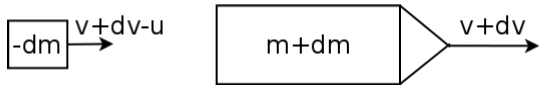


Figure 6.14: Rocket

of people with a rocket that possesses an initial mass m_i. Every instance, which is not necessarily regular, it releases an infinitesimal amount of mass at a velocity u with respect to itself, from its back. What is the speed of the rocket when its mass reaches m_f, assuming that its initial speed is v_i?

Let the mass of the rocket and its speed at an instance be m and v, respectively. When the rocket releases $-dm$ amount of mass ($-dm$ is positive as dm is negative), its mass becomes $m+dm$ and its velocity becomes $v+dv$ in the lab frame. Thus, the ejected fuel must have velocity $v+dv-u$ (Fig. 6.14).

By the conservation of momentum,

$$\begin{aligned}
mv &= (m+dm)(v+dv) - dm(v+dv-u) \\
dv &= -\frac{u}{m}dm, \\
v_f - v_i &= -u\int_{m_i}^{m_f} \frac{1}{m}dm = -u[\ln|m|]_{m_i}^{m_f} = u\ln\frac{m_i}{m_f} \\
v_f &= v_i + u\ln\frac{m_i}{m_f}.
\end{aligned}$$

The above example of rocket motion shows how to apply the conservation of momentum in solving problems. Unfortunately, the ln factor in the expression looks especially disheartening — particularly when this method of propulsion is the epitome of efficiency!

Systems under a Net External Force

The previous systems were not under the influence of a net external force — enabling the application of the conservation of momentum. For systems with variable masses under a net external force, we can apply the impulse-momentum theorem to the system, across an infinitesimal time interval dt.

Problem: A cart is traveling at a velocity v. Suppose that you begin to add sand traveling at a velocity u, to the cart at a mass per unit time σ, determine the force F you need to exert on the cart so that it travels at a constant velocity v.

Let the mass of the cart at the current instance be m. Then during a time interval dt, the cart gains an additional dm amount of sand, which was initially traveling at velocity u. Its final mass and velocity then become

$m + dm$ and $v + dv$ respectively. From the impulse-momentum theorem,

$$Fdt = \Delta p = (m + dm)(v + dv) - (mv + udm).$$

Fdt is the impulse delivered during the short time interval while the right-hand side represents the change in the combined momentum of the cart and the incoming sand. Discarding the second order term $dmdv$ and dividing the entire equation by dt,

$$F = \frac{dm}{dt}(v - u) + m\frac{dv}{dt}.$$

Since $\frac{dv}{dt} = 0$ and $\frac{dm}{dt} = \sigma$,

$$F = \sigma(v - u).$$

Problem: A chain of uniform linear mass density λ and length l is initially held motionless vertically downwards, with one of its ends just touching the table. It is then released. Find the normal force that the table exerts on the chain as a function of x, the distance that the top of the chain has fallen. Assume that when a part of the chain collides with the table, it goes to rest instantaneously.

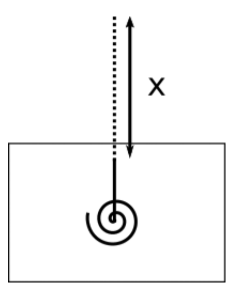


Figure 6.15: Falling chain

The normal force of the table on the chain has two roles. Firstly, it supports the weight of the chain that is already at rest on the table. Secondly, it has to stop the part of the falling chain that collides with the table. Let us analyze the force required for the second factor and then add the weight of the chain that already rests on the table to find the normal force. In a time interval dt, λvdt mass of chain that was initially traveling[4] at v comes

[4]The additional velocity of this falling part due to its acceleration by gravity is a second-order term and is thus negligible.

to rest. Thus, by the impulse-momentum theorem,

$$Fdt = vdm = \lambda v^2 dt.$$

v can be expressed in terms of x from the conservation of energy (or from the kinematics equations) since the moving part of the chain must be in free-fall due to there being no tension in the chain (see Chapter 4).

$$v^2 = 2gx$$
$$\implies F = 2\lambda gx.$$

Lastly, we can add back the weight of the motionless part of the chain to find the normal force,

$$N = F + \lambda gx = 3\lambda gx.$$

Through this approach, the reason why N abruptly plummets from $3\lambda gl$ to λgl when $x = l$ becomes lucid — there is no longer any falling segment that collides with the scale (and this component contributes $2\lambda gl$ to N when $x = l$).

Sometimes, if the change in mass can be expressed in terms of some distance travelled, the work-energy theorem can be applied over an infinitesimal distance — as illustrated in the following example.

Problem: A massless bucket initially contains a mass M of sand and is stationary at the origin. You then pull it in the positive x-direction via a constant tension T across the frictionless ground. If the bucket leaks sand at a rate $\frac{dm}{dx} = -\frac{M}{L}$ where m is the instantaneous mass of the bucket, determine its kinetic energy as a function of its x-coordinate x for $x < L$. ("An Introduction to Classical Mechanics")

Let the instantaneous kinetic energy of the bucket-cum-sand be $E(x)$. Consider the change in kinetic energy dE as the bucket travels an infinitesimal distance dx. By the work-energy theorem, the bucket gains Tdx amount of kinetic energy due to the work done on it. However, it also gains $\frac{dm}{m}E$ amount of kinetic energy (this quantity is negative as dm is negative) as it gains mass dm. Thus,

$$dE = Tdx + \frac{dm}{m}E.$$

Dividing the entire equation by dx,

$$\frac{dE}{dx} = T + \frac{dm}{dx} \cdot \frac{E}{m}.$$

Since $\frac{dm}{dx} \cdot \frac{E}{m} = -\frac{M}{L} \cdot \frac{E}{M - \frac{M}{L}x} = -\frac{E}{L-x}$,

$$\frac{dE}{dx} + \frac{E}{L-x} = T.$$

Multiplying the above by the integrating factor $\frac{1}{L-x}$,

$$\frac{1}{L-x}\frac{dE}{dx} + \frac{E}{(L-x)^2} = \frac{d\left(\frac{E}{L-x}\right)}{dx} = \frac{T}{L-x}$$

$$\int_0^{\frac{E}{L-x}} d\left(\frac{E}{L-x}\right) = \int_0^x \frac{T}{L-x}dx.$$

Simplifying,

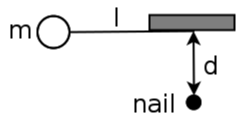
$$E = T(L-x)\ln\frac{L}{L-x}.$$

Problems

Conservation Laws

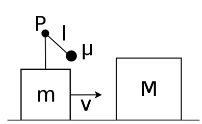
1. *Nail**

A pendulum with a mass m attached at its end is released at rest from a horizontal initial position. It then collides with a nail that is situated at a distance d below the top of the pendulum. Find minimum distance d in terms of l such that the mass will exhibit circular motion around the nail after the collision.



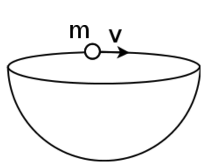
2. *Colliding Carts**

A cart of mass m and initial velocity v collides with another cart of mass M on a frictionless, horizontal ground and sticks to it. There is a small pendulum of length l with a mass μ attached to it, that is initially aligned with the vertical. Assuming that $\mu \ll m$, find the minimum initial velocity v required for the pendulum to exhibit circular motion around point P.



3. *Particle on Hemisphere**

A particle of mass m travels at an initial velocity v azimuthally on the top of a smooth hemisphere of radius r. Find the maximum azimuthal velocity of the particle in its motion afterwards, assuming that the hemisphere does not move or rotate.



4. *Raising a Pendulum**

A pendulum bob of mass m is attached to the ceiling via a massless string of length l that currently makes an angle θ_0 with respect to the vertical. If the bob and the string are in the x-y plane and the bob is at rest, determine the minimum velocity in the z-direction, u, that is required to be imparted to the bob for it to touch the ceiling.

5. *Collision with Spring**

A particle of mass m is traveling at velocity u on a frictionless, horizontal table. It then collides elastically with and sticks to a massless spring of spring constant k, which is attached to a block of mass M. Determine the maximum compression and extension of the spring in the motion afterwards.

6. *Maximum Height**

A block of mass m initially rests on a pair of frictionless rails with a mass M hanging vertically from it via a massless string of length l. Now, a bullet of mass m, traveling at velocity u parallel to the rails, is embedded into the block. Assuming that the string remains taut throughout the entire process, determine the maximum height that the mass M reaches above its original vertical position. How much work has been done by the tension in the string on mass M when M first reaches this maximum height?

7. *Crawling Ant**

An ant of mass m travels tangentially along a ring with radius r, mass M and negligible thickness. When it reaches the opposite side of the ring (you can mark a dot opposite to its initial position to indicate its destination), find the angle through which the ring has rotated. The ring can only rotate about its center and cannot translate.

8. *Ladder on Wall***

A ladder, of mass m and length $2l$, is initially held motionless along a wall at an angle θ_0 with respect to the wall. It is then released and the top end begins to slide down the wall whereas the bottom end slides along the ground. Assuming that all surfaces are frictionless, find $\ddot{\theta}$ as a function of θ by considering the total energy of the system and taking $\frac{dE}{dt} = 0$.

9. *Toppling Cube***

In Cartesian coordinates, the eight vertices of a cube lie at coordinates $(0,0,0)$, $(l,0,0)$, $(0,l,0)$, $(l,l,0)$, $(0,0,l)$, $(l,0,l)$, $(0,l,l)$ and (l,l,l). Determine the largest impulse J that can be delivered to the cube at point $(\frac{l}{2},0,h)$

such that the cube does not topple under the following conditions: (1) the edge of the cube at $y = l$, $z = 0$ is fixed, and (2) the ground ($z = 0$) on which the cube lies is frictionless.

10. *Spinning Earth***

Model the Earth as a uniform sphere of mass M and radius R, with its rotational axis defined to be the z-axis. Currently, a point particle of mass $m \ll M$ rests at one of the Earth's poles along the z-axis and the angular velocity of the Earth is ω_0. If the particle begins to move along a great circle on the surface of the Earth such that its angular coordinate from the z-axis is $\theta = \alpha t$ (where α is a constant and t is the time elapsed) determine the angle that the Earth has rotated by the time the particle reaches the opposite pole. Assume that the initial angular momentum of the Earth is large enough such that its angular velocity vector always lies along the z-axis.

11. *Connected Masses***

Two particles, of masses m and $2m$, lie on a frictionless, horizontal table along the x-axis in the xy-plane. A spring of spring constant k and zero relaxed length is attached to them. If the initial distance between them is l and m and $2m$ travel at v_1 and v_2 in the positive y-direction respectively, determine the minimum and maximum distances between them in their subsequent motion.

12. *Rolling over a Step****

A sphere of mass m and radius r approaches a rough step with height $h < r$. If the initial velocity of its center of mass is v and it rolls without slipping, under what conditions will the sphere collide with the step, rotate about the point of collision and move up the step?

Collisions

13. *Max Deflection**

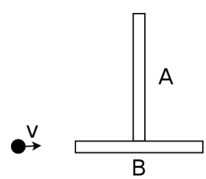
A particle of mass m_1 travels at speed v_1 in the positive x-direction and collides with a particle of mass $m_2 = \frac{1}{2}m_1$ which was travelling at speed v_2 in the positive y-direction. The maximum observed angle of deflection of m_1, after myriad experiments, is $\frac{\pi}{2}$ radians. Given m_1, m_2 and v_1, determine v_2.

14. *Particle Rocket***

n small particles are stacked on top of each other, with an infinitesimal distance between adjacent particles. Number the particles from 1 to n in a bottom-up fashion such that the ith particle has mass $f^{i-1}m$ where m is the mass of the first particle and f is a constant. If the entire array of particles is dropped from a height h onto horizontal ground, determine the velocity of the nth particle immediately after it has collided with the $(n-1)$th particle. Assume that all collisions are elastic.

15. *T-Shape***

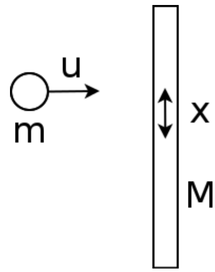
A T-shaped structure is formed from two uniform rods, A and B, each of length l, mass m and negligible thickness (see figure below). The midpoint of rod B is attached to one end of rod A. The structure is then put on a frictionless horizontal table. A point mass m, moving horizontally to the right at right angles to rod A (parallel to rod B), strikes the end of rod B with an initial velocity v and sticks to it.



(a) Write down the equation for conservation of linear momentum. Is energy conserved during the collision? Explain.
(b) Comment on the direction of the center of mass velocity of the combined system in the motion thereafter. Which point (possibly external to the body) has a constant linear velocity after the collision? Find the exact location of the point.
(c) Find the moment of inertia of the whole system (after the collision) about the point you found in (b).
(d) Solve for the angular velocity of the system after the collision.

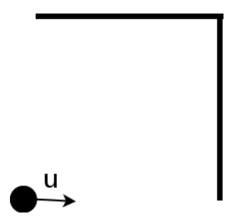
16. *Collision with Rod***

A ball of negligible size, mass m and initial speed u undergoes an elastic head-on collision with a uniform, stationary rod of mass M and length l, at a distance x above its center as shown in the figure below. If the final velocities of the ball and the center of the rod are aligned, determine their final velocities and the final angular speed ω of the rod. Determine the value(s) of x that maximise(s) ω. For this/these particular case(s), find the point on the rod that is instantaneously stationary right after the collision.



17. ***L-Shape Collision******

Two uniform rods, of mass M and length l each, are connected to form an "L-shape" as shown in the figure below. If the resultant structure is initially stationary and a particle of mass m undergoes a head-on elastic collision with one of its ends (such that the final velocity of the particle and the center of mass of the structure are directed along the same line), determine all values of $\frac{m}{M}$ such that a second collision occurs.



Work and Impulse

18. ***Zero Impulse****

A uniform rod of mass m and length l is pivoted at one of its ends and stands vertically. Determine the height h above the pivot at which an impulse should be delivered such that the impulse on the rod due to the pivot is zero.

19. ***Relativity of Work*****

An initially-stationary man of mass M throws a snowball of mass m at a relative velocity u away from him on frictionless ground. Determine the work done by the man's muscles (assuming that no heat is generated and that the mass of the man does not vary) in this process by applying the conservation of energy in the lab frame. Next, show that you obtain the same result if you directly compute the work done in the man's frame.

20. *Spinning Collision***

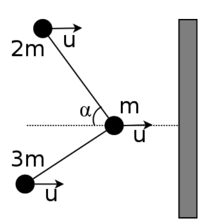
A sphere, of mass m and radius r, is spinning at an angular velocity ω_0 about an axis parallel to the plane of the table, with a coefficient of kinetic friction μ. If the sphere is released with zero initial translational velocity and collides with the table after its center of mass has fallen by a vertical height h, determine the angle θ between the vertical and the instantaneous velocity of the center of mass of the sphere after the collision. Assume that the center of mass of the sphere rebounds with the same vertical speed as that before the collision. Hint: There are two regimes of ω_0 to consider.

21. *Bouncing Mass***

A mass begins moving from a horizontal floor with a vertical velocity v_0 and horizontal velocity u_0. Every time it collides with the floor, its resultant vertical speed is a fraction e of that before the collision. If the coefficient of kinetic friction between the ground and the mass is μ, and the horizontal velocity of the mass becomes zero after exactly n collisions, determine n.

22. *Three Masses***

Three masses m, $2m$ and $3m$ are connected by two massless and rigid rods of length l which are currently perpendicular to each other, as shown in the figure. If the masses initially travel at velocity u towards a vertical wall and mass m undergoes a collision with the wall, determine the impulse delivered by the wall to mass m if the final horizontal velocity of mass m is zero. There is no friction between the wall and mass m. Assume that the tensions in the rods are strictly longitudinal (because they are massless).



23. *N Disks****

There are N disks, numbered from 1 to N, on a horizontal table. The ith disk has radius $f^{i-1}R$ and has initial angular velocity ω_i. If the disks are now arranged next to each other such that each disk touches disks of numbers adjacent to its own, determine the final angular velocity of the disk numbered 1. There is friction between the disks.

24. *Rebounding Mass****

A mass is released on a rough and massive inclined plane, at a distance l along the ramp, as measured from the bottom. The plane has an angle of inclination θ and there is a coefficient of kinetic friction μ between the mass and the plane. The bottom of the ramp is blocked by a massive barricade. The head-on collision between the mass and the blockade is governed by the coefficient of restitution e. Find the total distance traveled by the mass with the assumption that the coefficient of static friction is smaller than $\tan\theta$.

Variable Masses

25. *Drag Force on Sheet**

A massive sheet is surrounded by stationary sand of mass density ρ. If the sheet travels at a constant speed v, determine the drag force per unit area it experiences, assuming that collisions are elastic.

26. *Propelling a Car***

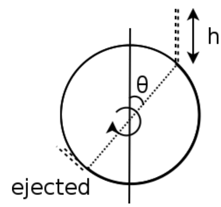
You begin to throw baseballs at speed u towards a car of mass M that is free to move frictionlessly on the ground. The baseballs leave your hand at a mass per unit time σ and bounce elastically off the car window, directly backwards. If the car starts at rest, find its speed and position as functions of time. ("An Introduction to Classical Mechanics")

27. *Drag Force on Sphere***

A massive sphere of radius R is surrounded by non-interacting air particles of mass density ρ. If the sphere travels at a speed v, determine the drag force it experiences, assuming collisions are elastic.

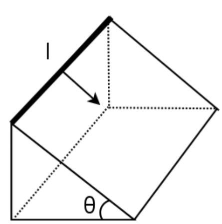
28. *Carrying Sand***

A massless circle of radius r is secured by a central axle that restricts it from translational motion and only permits rotational motion about its center. Initially-motionless sand is dropped from a height h at a mass rate σ and lands on the circle at a clockwise angular coordinate θ from the vertical. If the sand sticks to the circle and is carried to a clockwise angular coordinate θ from the bottom of the circle, where it is released (with no change in tangential velocity), determine the angular velocity ω of the circle when it has attained a steady state.



29. *Sweeping Duster****

A rod-shaped duster has length l and initial mass M. It is initially held motionless at the top of an inclined plane with an angle of inclination θ. The whole plane is covered with dust with a surface mass density σ. The duster is then given a slight push, picking up dust in its motion. Find its velocity v along the plane as a function of x, the distance along the surface of the inclined plane that it has traveled.



30. *Raindrop****

Model a raindrop as a blob that constantly maintains the shape of a homogeneous sphere with a constant density. Now consider a raindrop, initially of negligible size, that begins to fall through a uniform cloud of tiny water droplets. If the raindrop accumulates the water droplets, determine its acceleration.

Solutions

1. Nail*

The speed of the pendulum, v, immediately before the collision is given by the principle of the conservation of energy.

$$v = \sqrt{2gl}.$$

The angular momentum of the system is conserved about the nail during the collision. Thus, the pendulum mass still continues to travel at speed v after the collision (we can also argue that its horizontal momentum is conserved since tension can only be exerted along the rope, which is vertical at the point of collision). Notice that the pendulum is most likely to deviate from circular motion when it is directly above the nail as the radial component of its weight is the largest while the required centripetal force is the smallest — implying that the tension in the string is minimum. In the critical case where the mass is just able to exhibit circular motion, the weight of the bob solely provides the centripetal force at the top of the circle (the tension at this juncture is zero). If the speed of the bob at the top of the circle is v',

$$mg = \frac{mv'^2}{l-d}.$$

Furthermore, by the conservation of energy,

$$\frac{1}{2}mv^2 = \frac{1}{2}mv'^2 + 2mg(l-d)$$
$$d = \frac{3}{5}l.$$

2. Colliding Carts*

The total momentum of the two carts is conserved during the collision. Note that the pendulum mass still travels at v as there is no impulsive force on it to instantaneously change its velocity. Letting the final velocity of the two carts be v',

$$mv = (m+M)v'$$
$$v' = \frac{mv}{m+M}.$$

Consider the frame of the cart, which is approximately an inertial frame despite the tension that acts on it. This is because $m \gg \mu$, where the change in the momentum of the pendulum, and thus the momentum of the cart, is on the order of μ. The velocity of the pendulum in this frame immediately after the collision is

$$v_r = v - v'.$$

Assuming that it is able to exhibit circular motion, let the velocity of the mass at the top of the circle be v_r' in the frame of the cart. Then, the gravitational force and tension must provide the centripetal force.

$$\mu g + T = \frac{\mu v_r'^2}{l}.$$

In the boundary case, v_r' is just large enough that $T = 0$. When v_r' is too small, T is negative, which implies that the bob has deviated from circular motion beforehand. Thus in the limiting case,

$$v_r' = \sqrt{gl}.$$

By the conservation of energy,

$$\begin{aligned}\frac{1}{2}\mu v_r^2 &= \frac{1}{2}\mu v_r'^2 + 2\mu gl \\ v_r &= \sqrt{5gl} = v - v' \\ v - \frac{mv}{m+M} &= \sqrt{5gl} \\ v &= \frac{m+M}{M}\sqrt{5gl}.\end{aligned}$$

3. Particle on Hemisphere*

We see that the angular momentum of the particle about a vertical axis through the center of the hemisphere is conserved, as the normal and gravitational forces on the particle produce no torque along that direction. Furthermore, it is important to note that when the particle attains its maximum azimuthal velocity, its velocity in the polar direction is zero. This is because the particle reaches its maximum azimuthal velocity at its lowest height, as implied by the conservation of angular momentum. If it were to have a velocity in the polar direction at that instant, its height at the next instant will be lower — leading to a contradiction.[5] We let the particle be at angle θ from

[5] In fact, the particle will undergo circular motion at a constant height when it attains the minimum height. This is contrary to our common intuition which suggests that the particle should eventually drop to the bottom. However, this occurs only because of the friction acting on the particle, which produces a torque that reduces L_z in real life.

the vertical axis when it reaches its minimum height. Then, conservation of angular momentum along the vertical axis gives

$$L_z = mrv = mr\sin\theta v_{max}.$$

By the conservation of energy,

$$\frac{1}{2}mv^2 + mgr\cos\theta = \frac{1}{2}mv_{max}^2.$$

Solving,

$$v_{max} = \sqrt{\frac{1}{2}\left(v^2 + \sqrt{v^4 + 16g^2r^2}\right)}.$$

4. Raising a Pendulum*

The angular momentum of the bob, with respect to a vertical axis passing through the point on the ceiling at which the string is attached, must be conserved as there is no external torque in the vertical direction. Thus, let v be the final azimuthal velocity of the pendulum when it reaches the ceiling. Then

$$ml\sin\theta_0 u = mlv$$
$$v = u\sin\theta_0.$$

In order for u to be minimum, the velocity of the pendulum bob, perpendicular to the plane of the ceiling, must be zero when it reaches the ceiling (as energy must still be conserved). Applying the conservation of energy to the bob (as the tension in the inextensible string does no work on the bob),

$$\frac{1}{2}mu^2 = mgl\cos\theta_0 + \frac{1}{2}mv^2$$
$$u = \sqrt{\frac{2gl}{\cos\theta_0}}.$$

5. Collision with Spring*

When the spring is at its maximum compression or extension, the particle and the block must be traveling at the same velocity v, as a relative velocity would imply that the relative separation at an earlier or later instance in time will either be greater or smaller (the exact change depends on whether the particle is faster than the block or vice-versa). By conservation of momentum,

$$v = \frac{mu}{M+m}.$$

By conservation of energy,

$$\frac{1}{2}mu^2 = \frac{1}{2}(M+m)v^2 + \frac{1}{2}kx^2,$$

where x is the maximum extension (defined to be positive) or compression (defined to be negative) of the spring. Solving for x,

$$x = \pm\sqrt{\frac{mM}{k(m+M)}}u.$$

The positive value refers to a maximum extension of $\sqrt{\frac{mM}{k(m+M)}}u$ while the negative value indicates a maximum compression of $\sqrt{\frac{mM}{k(m+M)}}u$ as well.

6. Maximum Height*

When M attains its maximum height, its instantaneous vertical velocity must be zero. Since the vertical velocity of the block that it is connected to is also zero, their horizontal velocities must be identical for the string to remain taut (length of the string is preserved). This is the crucial observation. Following from this and the conservation of momentum, the velocities of the block and the mass at the required juncture are

$$v = \frac{mu}{2m+M}.$$

Furthermore, the energy of the system after the collision is conserved. The total kinetic energy of the system directly after the collision is due to that of a block with mass $2m$ travelling at velocity $\frac{u}{2}$ (owing to the conservation of momentum). Thus, the total initial kinetic energy is

$$\frac{1}{2}\cdot 2m \cdot \frac{u^2}{4} = \frac{mu^2}{4}.$$

Applying the conservation of energy,

$$\frac{1}{2}(2m+M)v^2 + Mg\Delta h = \frac{mu^2}{4}$$

$$\Delta h = \frac{mu^2}{4Mg} - \frac{m^2u^2}{2(2m+M)Mg},$$

where Δh is the change in height of M. The work done by tension on mass M up till this point results in the total change in its mechanical energy, which is

$$W_T = \frac{1}{2}Mv^2 + Mg\Delta h = \frac{mu^2}{4} - mv^2 = \left(\frac{m}{4} - \frac{m^3}{(2m+M)^2}\right)u^2.$$

Now, you may wonder why a tension in an inextensible string can lead to work done. Well, even though the length of the string does not change, the entire string is constantly displaced at a certain velocity — engendering work done as the point of application of its force is moving.

7. Crawling Ant*

The key point is to note that when the ant moves, the ring rotates in the opposite direction. Let v be the tangential velocity of the ant and ω be the angular velocity of the ring, defined to be positive in the opposite direction. Let the total time taken for the ant to reach the opposite side be τ. Then,

$$\int_0^\tau v dt + \int_0^\tau r\omega dt = \pi r.$$

The left term is the distance covered by the ant while the right term is the distance that a point on the ring has rotated across. Furthermore, the angular momentum of the ant-cum-ring system about the center of the ring is conserved as the forces between the ant and the ring act along the same line.

$$0 = mrv - Mr^2\omega,$$

where we have used the fact that the moment of inertia of a ring about its center is Mr^2. Then,

$$v = \frac{M}{m}r\omega,$$

$$\int_0^\tau \left(1 + \frac{M}{m}\right) r\omega dt = \pi r.$$

Thus, the angle that the ring has rotated is

$$\int_0^\tau \omega dt = \frac{\pi}{1 + \frac{M}{m}}.$$

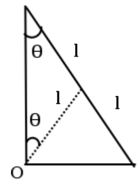


Figure 6.16: Ladder on wall

8. Ladder on Wall**

Referring to Fig. 6.16, the position and velocity of the center of mass are

$$\begin{aligned} x_{CM} &= l\sin\theta, \\ y_{CM} &= l\cos\theta, \\ \dot{x}_{CM} &= l\cos\theta\dot{\theta}, \\ \dot{y}_{CM} &= -l\sin\theta\dot{\theta}. \end{aligned}$$

The total mechanical energy of the ladder is

$$E = mgl\cos\theta + \frac{1}{2}m(\dot{x}_{CM}^2 + \dot{y}_{CM}^2) + \frac{1}{2}I_{CM}\dot{\theta}^2 = mgl\cos\theta + \frac{2}{3}ml^2\dot{\theta}^2,$$

where $I_{CM} = \frac{1}{12} \cdot m \cdot (2l)^2 = \frac{1}{3}ml^2$ for the ladder.

$$\frac{dE}{dt} = 0 \implies -mgl\sin\theta\dot{\theta} + \frac{4}{3}ml^2\dot{\theta}\ddot{\theta} = 0.$$

$$\ddot{\theta} = \frac{3g\sin\theta}{4l}.$$

9. Toppling Cube**

Note that if the cube is pivoted about the particular edge, there is generally an impulsive force on the cube by the pivot in addition to the impulse J. Thus, we cannot use J to determine the linear momentum of the cube. However, we can calculate the angular momentum of the cube about the fixed edge, as the impulse force due to the pivot does not produce any angular impulse about the pivot. By the angular impulse-momentum theorem,

$$L_{edge} = Jh = I_{edge}\omega,$$

where $I_{edge} = \frac{1}{6}ml^2 + \frac{1}{2}ml^2 = \frac{2}{3}ml^2$ for a cube. The total mechanical energy of the cube is thus

$$\begin{aligned} E &= \frac{1}{2} I_{edge}\omega^2 + \frac{mgl}{2} \\ &= \frac{1}{2}\frac{L_{edge}^2}{I_{edge}} + \frac{mgl}{2} \\ &= \frac{1}{2}\frac{J^2h^2}{\frac{2}{3}ml^2} + \frac{mgl}{2} \\ &= \frac{3J^2h^2}{4ml^2} + \frac{mgl}{2}. \end{aligned}$$

The cube will topple if it has a non-zero angular velocity when its center of mass reaches its maximum height as the torque thereafter will cause its angular velocity to accelerate further, in the direction that causes it to topple. Thus, in the boundary case, the total mechanical energy is just sufficient to raise the center of mass of the cube to this height.

$$\begin{aligned} \frac{3J^2h^2}{4ml^2} + \frac{mgl}{2} &= \frac{\sqrt{2}mgl}{2} \\ J &= \sqrt{\frac{2m^2gl^3}{3h^2}(\sqrt{2}-1)}. \end{aligned}$$

This is the maximum value of J for which the cube will not topple. In the case where the ground is frictionless and the cube is not pivoted, the only impulse is J. Applying the angular impulse-momentum theorem, the angular momentum about the center of mass of the cube is

$$L_{CM} = J\left(h - \frac{l}{2}\right) = I_{CM}\omega,$$

where ω is the initial angular velocity of the cube. The total mechanical energy is

$$\begin{aligned} E &= \frac{1}{2}mv_{CM}^2 + \frac{1}{2}I_{CM}\omega^2 + \frac{mgl}{2} \\ &= \frac{p^2}{2m} + \frac{L_{CM}^2}{2I_{CM}} + \frac{mgl}{2} \end{aligned}$$

$$= \frac{J^2}{2m} + \frac{J^2\left(h - \frac{l}{2}\right)^2}{2 \cdot \frac{1}{6}ml^2} + \frac{mgl}{2}$$

$$= \frac{J^2}{2m}\left(1 + \frac{6\left(h - \frac{l}{2}\right)^2}{l^2}\right) + \frac{mgl}{2}.$$

In the same boundary case,

$$E = \frac{\sqrt{2}mgl}{2}$$

$$\implies J = \sqrt{\frac{m^2gl^3}{l^2 + 6\left(h - \frac{l}{2}\right)^2}(\sqrt{2} - 1)}.$$

10. Spinning Earth**

The angular momentum of the Earth-and-particle system about the z-axis must be conserved due to the lack of net external torques. The initial angular momentum is $\frac{2}{5}MR^2\omega_0$. When the particle is at coordinate $\theta = \alpha t$, the moment of inertia of the combined system is $\frac{2}{5}MR^2 + mR^2\sin^2\alpha t$. Thus, the angular velocity of the Earth-and-particle system about the z-direction is

$$\omega = \frac{\frac{2}{5}MR^2}{\frac{2}{5}MR^2 + mR^2\sin^2\alpha t}\omega_0$$

$$= \frac{\omega_0}{1 + \frac{5m}{2M}\sin^2\alpha t}$$

$$\approx \omega_0\left(1 - \frac{5m}{2M}\sin^2\alpha t\right).$$

Note that the angular velocities of the Earth and particle must be identical for the particle to travel along the same great circle. Now, the total angle ϕ that the Earth has rotated when the particle reaches $\theta = \pi$ is obtained by integrating the above expression with respect to t.

$$\phi = \int_0^{\frac{\pi}{\alpha}} \omega dt$$

$$= \int_0^{\frac{\pi}{\alpha}} \left(\omega_0 - \frac{5m}{2M}\omega_0\sin^2\alpha t\right) dt$$

$$= \frac{\pi\omega_0}{\alpha}\left(1 - \frac{5m}{4M}\right).$$

11. Connected Masses**

In the center of mass frame which travels at

$$v_{CM} = \frac{v_1 + 2v_2}{3}$$

in the positive y-direction, the velocities of the two particles in this frame are

$$v_1' = \frac{2v_1 - 2v_2}{3},$$
$$v_2' = \frac{v_2 - v_1}{3}.$$

In this center of mass frame, the center of mass, which lies at a distance $\frac{l}{3}$ away from $2m$, must not shift, due to the lack of a net external force on this system. Therefore, these two particles must orbit about the center of mass at some common angular velocity ω when they are at their closest or furthest approach. ω is defined to be positive in the direction of the initial angular momentum of m about the center of mass. Let an extremal distance (maximum or minimum) between the particles be $3r$, split into $2r$ and r between the center of mass and the particles. As there is no net external torque, we can apply the conservation of angular momentum about the center of mass to get

$$\frac{2l}{3} \cdot m \cdot v_1' - \frac{l}{3} \cdot 2m \cdot v_2' = 2mr^2\omega + m(2r)^2\omega$$
$$\omega = \frac{l(v_1 - v_2)}{9r^2}.$$

Furthermore, at the closest or furthest distance of approach, the radial velocities of masses must be zero. Applying the conservation of energy,

$$\frac{1}{2}m\left(\frac{2v_1 - 2v_2}{3}\right)^2 + \frac{1}{2} \cdot 2m \cdot \left(\frac{v_2 - v_1}{3}\right)^2 + \frac{1}{2}kl^2$$
$$= \frac{1}{2} \cdot m \cdot (2r)^2\omega^2 + \frac{1}{2} \cdot 2m \cdot r^2\omega^2 + \frac{1}{2}k \cdot (3r)^2 = 0.$$

Substituting ω and simplifying,

$$r^4 - \left(\frac{2m}{27k}(v_1 - v_2)^2 + \frac{l^2}{9}\right) r^2 + \frac{2ml^2}{243k}(v_1 - v_2)^2 = 0.$$

Without attempting to solve the above equation, we can already state a root for r^2. Since the initial state of the system has no relative radial velocity,

$r^2 = \frac{l^2}{9}$ must be a solution to the above equation. This enables us to factorize the above into

$$\left(r^2 - \frac{2m}{27k}(v_1 - v_2)^2\right)\left(r^2 - \frac{l^2}{9}\right) = 0.$$

The other possible expression for r is

$$r = \sqrt{\frac{2m}{27k}}|v_1 - v_2|$$
$$\implies 3r = l \quad \text{or} \quad \sqrt{\frac{2m}{3k}}|v_1 - v_2|.$$

The maximum separation is the larger of the two; vice-versa for the minimum separation.

12. Rolling over a Step***

In this problem, there are actually two different stages of motion. Firstly, the sphere collides with the step — a process in which the linear momentum and energy of the sphere are not conserved. Afterwards, the sphere might rotate about the point of contact with the step or rebound from it.

Though the linear momentum and energy of the sphere are not conserved during the collision with the step, if we choose our stationary origin O to be at the point of contact with the step, the angular momentum about O is conserved. This is because the only impulsive force is the contact force on the sphere due to the step. This force acts at the origin, producing no torque. The torque due to the weight of the sphere with respect to O imparts negligible angular impulse during the short collision period. Let the initial angular velocity of the sphere about its center be ω, and the angular velocity of the sphere about O be ω' after the collision. By the conservation of angular momentum,

$$mv(r-h) + I_{CM}\omega = (I_{CM} + mr^2)\omega',$$

where $I_{CM} = \frac{2}{5}mr^2$, the moment of inertia of the sphere about an axis through its center. We have used the parallel axis theorem to calculate the moment of inertia of the sphere about O on the right-hand side of the equation. Moreover, recall that we also have the initial non-slip condition.

$$v = r\omega.$$

Let us assume that the sphere manages to rotate about O. During this rotational motion, the total mechanical energy of the system is conserved.

Furthermore, in the boundary case where the sphere just manages to roll up the step, its final angular velocity at the top of the step is zero. Thus, in that boundary case,

$$\frac{1}{2}\cdot(I+mr^2)\omega'^2 = mgh$$

$$\frac{7}{10}mr^2\omega'^2 = mgh,$$

as the vertical coordinate of the center of mass increases by h. Solving the equations,

$$v - \frac{r}{7r-5h}\sqrt{70gh}.$$

Thus for the sphere to roll up,

$$v \geq \frac{r}{7r-5h}\sqrt{70gh}.$$

This is the condition on v for the sphere to roll up the step, assuming that it is able to rotate about O (i.e. does not lose contact).

Finally, the sphere might not be able to roll up the step if it loses contact with point O. After it stops rotating about O, the sphere will just fall back down due to its own weight. Thus, we must first determine where the sphere is most likely to lose contact with the step. This is in fact at the bottom of the step when the sphere initially begins to rotate about O. The angular velocity of the sphere at this point is the largest (so the required centripetal force follows suit) and the component of the gravitational force in the radial direction towards O is the smallest — signifying that the normal force N exerted on the sphere by the step is the smallest if it is indeed able to rotate about O. Analyzing the forces in the radial direction when the sphere is at the bottom of the step,

$$mg\sin\theta - N = mr\omega'^2,$$

where $\theta = \sin^{-1}\frac{r-h}{r}$ is the angle subtended by a line joining O and the center of the sphere and the horizontal. The boundary case where $N = 0$ occurs when

$$\omega' = \frac{\sqrt{g(r-h)}}{r},$$

which requires

$$v = \frac{7r}{7r-5h}\sqrt{g(r-h)}.$$

Combining the two conditions,

$$\frac{r}{7r-5h}\sqrt{70gh} \leq v \leq \frac{7r}{7r-5h}\sqrt{g(r-h)}.$$

Dividing the upper bound by the lower bound, it is also necessary that

$$r \geq \frac{119}{49}h.$$

13. Max Deflection*

Considering the initial frame of m_2, m_1 travels at an angle $\tan^{-1}\frac{v_2}{v_1}$ with respect to the horizontal. It was previously derived that the maximum angle of deflection for a particle of mass m_1 colliding with a stationary particle of mass $m_2 < m_1$ is $\sin^{-1}\frac{m_2}{m_1}$. Thus, the maximum angle between m_1's final velocity and the x-axis in the frame of m_2 is the sum of this angle and the angle that m_1 makes in m_2's frame.

$$\theta_{max} = \tan^{-1}\frac{v_2}{v_1} + \sin^{-1}\frac{m_2}{m_1} = \tan^{-1}\frac{v_2}{v_1} + \frac{\pi}{6}.$$

Since the maximum deflection angle is $\frac{\pi}{2}$ in the lab frame and the initial frame of m_2 only traveled in the y-direction, $\theta_{max} = \frac{\pi}{2}$.

$$\tan^{-1}\frac{v_2}{v_1} = \frac{\pi}{3}$$
$$v_2 = \sqrt{3}v_1.$$

14. Particle Rocket**

Before any collision event occurs, the particles all essentially travel at speed $\sqrt{2gh}$ downwards (as their sizes are negligible), by the conservation of energy. Let the velocity of the ith particle, immediately after it has collided with the $(i-1)$th particle be v_i (positive upwards). We have derived that in a general one-dimensional elastic collision between two particles m_A and m_B with initial velocities u_A and u_B, the final velocity of m_B is

$$v_B = \frac{2m_Au_A + (m_B - m_A)u_B}{m_A + m_B}.$$

The equation above can be applied to the collision between the ith particle and the $(i-1)$th particle by setting $m_A = m_{i-1} = f^{i-2}m$ and $m_B = m_i = f^{i-1}m$ for $i \geq 2$. The initial velocities in this case are $u_A = v_{i-1}$ and

$u_B = -\sqrt{2gh}$ (negative as m_i is still traveling downwards).

$$v_i = \frac{2f^{i-2}mv_{i-1} - \left(f^{i-1} - f^{i-2}\right) m\sqrt{2gh}}{\left(f^{i-1} + f^{i-2}\right) m}$$

$$v_i = \frac{2}{f+1}v_{i-1} - \frac{f-1}{f+1}\sqrt{2gh}$$

$$v_i + \sqrt{2gh} = \frac{2}{f+1}\left(v_{i-1} + \sqrt{2gh}\right).$$

Since the base case is $v_1 = \sqrt{2gh}$ (as the bottom-most particle is reflected from the ground),

$$v_n + \sqrt{2gh} = \left(\frac{2}{f+1}\right)^{n-1}\left(v_1 + \sqrt{2gh}\right)$$

$$v_n = \sqrt{2gh}\left(\frac{2^n}{(f+1)^{n-1}} - 1\right).$$

15. T-Shape**

(a) Let v' denote the final velocity of the combined center of mass of the "T-shape" and the particle m. By the conservation of momentum,

$$mv = 3mv'$$

$$v' = \frac{v}{3}.$$

Energy is not conserved due to the inelastic collision between the particle and rod B which generates heat and sound energies that are dissipated to the external environment. An obvious way of seeing this is to consider the center of mass frame which travels at v' with respect to the lab frame. Originally, the particle m travels at $\frac{2v}{3}$ while the "T-shape" purely translates at a center of mass velocity $-\frac{v}{3}$. After the collision, the combined system rotates about its combined center of mass but does not translate. Choosing an origin at the same vertical level as the center of mass of the "T-shape" (not including the particle) and applying the conservation of angular momentum, the final angular momentum of the combined system about its center of mass (since there is no translational component of angular momentum) is identical to the initial angular momentum of the particle about the origin. However, since the moment of inertia of the combined system about its center of mass is larger than that[6] of the particle, the final rotational kinetic energy

[6] The instantaneous moment of inertia of the particle can be computed by taking mass m multiplied by the perpendicular distance between the origin and the particle at the juncture of collision.

of the combined system (recall that there should not be any translational component of energy) is smaller than the initial energy of the particle in this new frame — signifying energy loss.

(b) The direction of the velocity of the center of mass of the combined system is still rightwards by the conservation of momentum. The center of mass of the combined system has a constant linear velocity after the collision as no net external force acts on the combined system. The center of mass of the "T-shape" is $\frac{l}{4}$ above the connection point — implying that the center of mass of the combined system is $\frac{l}{6}$ left of the connection point and $\frac{l}{4} \cdot \frac{2}{3} = \frac{l}{6}$ above the connection point.

(c) The moment of inertia of the particle about the center of mass is $m\left(\frac{l^2}{3^2} + \frac{l^2}{6^2}\right) = \frac{5}{36}ml^2$. On the other hand, the moment of inertia of rod B about the same point is $\frac{1}{12}ml^2 + m\left(\frac{l^2}{6^2} + \frac{l^2}{6^2}\right) = \frac{5}{36}ml^2$ while that of rod A is $\frac{1}{12}ml^2 + m\left(\frac{l^2}{3^2} + \frac{l^2}{6^2}\right) = \frac{8}{36}ml^2$ by the parallel axis theorem. The total moment of inertia about the center of mass is obtained from summing the individual contributions.

$$I_{CM} = \frac{1}{2}ml^2.$$

(d) Applying the conservation of angular momentum with respect to the center of mass of the combined system,

$$m \cdot \frac{l}{6} \cdot v = I_{CM}\omega$$
$$\omega = \frac{v}{3l},$$

in the anti-clockwise direction.

16. Collision with Rod**

Let the final velocities of the ball and the center of the rod be v_1 and v_2 respectively. Keeping in mind that the moment of inertia of the uniform rod about its center is $\frac{1}{12}Ml^2$, the conservations of momentum and energy yield

$$mu = mv_1 + Mv_2 \implies m(u - v_1) = Mv_2,$$
$$\frac{1}{2}mu^2 = \frac{1}{2}mv_1^2 + \frac{1}{2}Mv_2^2 + \frac{1}{24}Ml^2\omega^2.$$

Applying the conservation of angular momentum (while taking ω to be positive clockwise) about an origin, that passes through the line depicting the initial velocity of the ball,

$$\frac{1}{12}Ml^2\omega - Mxv_2 = 0$$

$$\omega = \frac{12xv_2}{l^2}.$$

Substituting this expression for ω into the second equation,

$$m(u - v_1)(u + v_1) = Mv_2^2\left(1 + \frac{12x^2}{l^2}\right).$$

Dividing the first equation squared by this equation,

$$\frac{m(u - v_1)}{u + v_1} = \frac{M}{1 + \frac{12x^2}{l^2}}.$$

Solving,

$$v_1 = \frac{1-k}{1+k}u,$$

where $k = \frac{M}{m\left(1+\frac{12x^2}{l^2}\right)}$. Substituting this expression for v_1 into the conservation of momentum equation,

$$v_2 = \frac{m(u - v_1)}{M} = \frac{2mk}{M(1+k)}u$$

$$\implies \omega = \frac{24mkx}{Ml^2(1+k)}u = \frac{24xu}{\left(1 + \frac{M}{m} + \frac{12x^2}{l^2}\right)l^2}.$$

To maximize ω, one has to find $\frac{d\omega}{dx} = 0$, which requires

$$\frac{\left(1 + \frac{M}{m} + \frac{12x^2}{l^2}\right) - \frac{24x^2}{l^2}}{\left(1 + \frac{M}{m} + \frac{12x^2}{l^2}\right)^2} = 0$$

$$\implies x^2 = \frac{m+M}{12m}l^2$$

$$x = \pm\sqrt{\frac{m+M}{12m}}l.$$

Both values of x above indeed correspond to maxima in terms of angular speed (i.e. $|\omega|$) as one can easily check by finding the first derivative evaluated at adjacent values. The positive result corresponds to the largest clockwise angular velocity while the other represents the largest anti-clockwise angular velocity. Note that the above expression for x is only valid for $|x| \leq \frac{l}{2}$, or equivalently, $M \leq 2m$. For $M > 2m$, $\frac{d\omega}{dx} > 0$ for $0 \leq x \leq \frac{l}{2}$ and $\frac{d\omega}{dx} < 0$ for $-\frac{l}{2} \leq x \leq 0$, such that the rod should be hit at its ends $x = \pm\frac{l}{2}$ to maximise $|\omega|$. It is interesting to note that the values of x that maximise the angular speed do not always correspond to the ends of the rod (see Ref. [2]). Moving on, the next problem is essentially about locating the instantaneous center of rotation (ICoR) right after the collision. The distance between the center of the rod and the ICoR is

$$\left|\frac{v_2}{\omega}\right| = \left|\frac{l^2}{12x}\right| = \begin{cases} \sqrt{\frac{m}{12(m+M)}}l & \text{for } M \leq 2m \\ \frac{l}{6} & \text{for } M > 2m. \end{cases}$$

Whether the ICoR is located above or below the center depends on the direction of angular velocity. An anti-clockwise angular velocity corresponds to an ICoR above the center and vice-versa for a clockwise angular velocity.

17. L-Shape Collision***

We first compute the moment of inertia of the "L-shape" structure with respect to its center of mass. Its center of mass is located at a vertical distance $\frac{l}{4}$ and horizontal distance $\frac{l}{4}$ away from the point of connection. Thus, by applying the parallel axis theorem,

$$I = 2 \cdot \left(\frac{1}{12}Ml^2 + \frac{1}{8}Ml^2\right) = \frac{5}{12}Ml^2.$$

Let v_1, v_2 and ω be the final velocities of the particle and the center of mass of the structure and the final angular velocity (positive anti-clockwise) of the structure respectively. By the conservation of momentum,

$$m(u - v_1) = 2Mv_2.$$

By the conservation of energy,

$$\frac{1}{2}mu^2 = \frac{1}{2}mv_1^2 + Mv_2^2 + \frac{1}{2}I\omega^2.$$

Applying the conservation of angular momentum about a point on the line that depicts the initial velocity of the particle (taking anti-clockwise to be positive),

$$I\omega - 2M \cdot \frac{3l}{4}v_2 = 0,$$

as the original angular momentum about the same origin was zero.

$$\implies \omega = \frac{18v_2}{5l}.$$

Substituting this into the third equation,

$$\frac{1}{2}m(u - v_1)(u + v_1) = \frac{37}{10}Mv_2^2.$$

Dividing this equation by the square of the second equation and simplifying,

$$v_1 = \frac{37\varepsilon - 20}{37\varepsilon + 20}u,$$

where $\varepsilon = \frac{m}{M}$. Solving for the other variables,

$$v_2 = \frac{20\varepsilon}{37\varepsilon + 20}u,$$
$$\omega = \frac{72\varepsilon}{37\varepsilon + 20} \cdot \frac{u}{l}.$$

There are a few cases under which a second collision can occur. The first is if $v_1 = v_2$ such that the structure collides with the particle after one complete rotation. This requires

$$\frac{37\varepsilon - 20}{37\varepsilon + 20}u = \frac{20\varepsilon}{37\varepsilon + 20}u$$
$$\implies \varepsilon = \frac{m}{M} = \frac{20}{17}.$$

Now, let us consider other cases. If $v_1 > v_2$, the particle will never collide with the structure again. Otherwise if $v_1 < v_2$, the only possible collision configuration occurs when the structure has rotated $\frac{\pi}{2} + n2\pi$ radians where $n \in \mathbb{Z}$, $n \geq 0$. Then, the center of mass of the structure must have traveled a distance $\frac{l}{2}$ relative to the particle at this juncture, as depicted in the figure below.

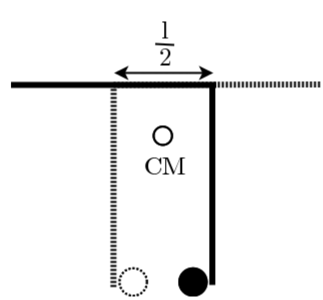


Figure 6.17: First and second (in dotted lines) collisions

Note that $\frac{l}{2}$ stems from the fact that the distance between the vertical rods in the original configuration and that after a $\frac{\pi}{2}$-radian rotation about its center of mass is $\frac{l}{2}$. Thus,

$$\frac{l}{2(v_2 - v_1)} = \frac{T}{4} + nT,$$

where T is the period of the structure's rotation. Rearranging,

$$\frac{l}{2T} = \left(\frac{1}{4} + n\right)(v_2 - v_1).$$

Substituting $\frac{1}{T} = \frac{\omega}{2\pi}$ and simplifying yields

$$\varepsilon = \frac{5 + 20n}{\frac{18}{\pi} + \frac{17}{4} + 17n}.$$

18. Zero Impulse*

Let the impulse delivered to the rod be J. Then, the angular momentum of the rod about the pivot is Jh after the impulse has been delivered, as the angular impulse imparted by the possible impulsive contact force exerted by the pivot is zero with respect to the pivot.

$$\frac{1}{3}ml^2\omega = Jh,$$

where ω is the instantaneous angular velocity of the rod. If the impulse on the rod due to the pivot is indeed zero, the total impulse delivered to the rod must be J. Thus,

$$J = mv_{CM} = \frac{ml\omega}{2}.$$

Solving the two equations above,

$$h = \frac{2}{3}l.$$

19. Relativity of Work**

Define the x-axis to be along the relevant direction in this one-dimensional motion. Suppose that the final velocity of the man is $u - v_0$ while that of the snowball is $-v_0$. Since the initial momentum of the combined system is zero, by the conservation of momentum,

$$\begin{aligned} M(u - v_0) - mv_0 &= 0 \\ v_0 &= \frac{Mu}{m+M} \\ u - v_0 &= \frac{mu}{m+M}. \end{aligned}$$

The work done by the man's muscles is the total increase in mechanical energy of the combined system.

$$W = \frac{1}{2}mv_0^2 + \frac{1}{2}M(u - v_0)^2 = \frac{mM}{2(m+M)}u^2.$$

The other perspective is significantly harder. Define the instantaneous velocity of the man in the lab frame as V and that of the snowball as $-v$ when they have yet to reach their final values. Let the instantaneous force exerted by the man on the snowball be F in the negative x-direction. Note that the rate of work done by the man in his own frame is not FV or Fv but rather $F(v+V)$ (force multiplied by the relative velocity between the man and the snowball) as the distance covered by the snowball in the man's frame (perhaps, due to the extension of the man's arm) increases at the rate $v_{rel} = v + V$. By the conservation of momentum in the lab frame,

$$MV - mv = 0 \implies V = \frac{mv}{M}$$
$$v_{rel} = \left(1 + \frac{m}{M}\right)v.$$

The rate of work done by the man in his own frame is thus

$$\frac{dW}{dt} = Fv_{rel} = \left(1 + \frac{m}{M}\right)mav,$$

where a is the acceleration of the snowball in the negative x-direction, in the lab frame. The total work done is then

$$W = \int_0^t \left(1 + \frac{m}{M}\right)mavdt = \int_0^{v_0^2} \frac{1}{2}\left(1 + \frac{m}{M}\right)md(v^2) = \frac{mM}{2(m+M)}u^2.$$

20. Spinning Collision**

The sphere may have stopped slipping during the collision or continues to slip after the collision, depending on the magnitude of ω_0. To determine the conditions under which the sphere continues slipping, consider the maximum impulse and angular impulse delivered by friction. During the collision, the impulse delivered by the normal force is

$$J_N = 2mv,$$

where v is the speed of the center of the sphere before the collision.

$$v = \sqrt{2gh}.$$

The maximum impulse delivered by friction is thus

$$J_f = 2\mu mv.$$

The maximum angular impulse delivered by friction about the center is then

$$I_f = -2\mu mvr,$$

where the negative sign reflects the fact that the angular impulse is opposite in direction to ω_0. The maximum horizontal velocity of the center of mass and the minimum angular velocity of the sphere after the collision are then

$$u_x = \frac{J_f}{m} = 2\mu v,$$
$$\omega = \omega_0 + \frac{I_f}{I} = \omega_0 - \frac{5\mu v}{r}.$$

In order for the sphere to continue slipping,

$$r\omega > u_x$$
$$\implies \omega_0 > \frac{7\mu v}{r}.$$

In this regime, the horizontal velocity of the center of mass is $2\mu v$. Thus,

$$\theta = \tan^{-1}\frac{u_x}{v} = \tan^{-1} 2\mu.$$

When $\omega_0 \leq \frac{7\mu v}{r}$, the sphere stops slipping during the collision. Thus, the maximum impulse I_f is not completely delivered and we cannot conclude that $u_x = 2\mu v$. However, we now know that

$$u_x = r\omega,$$

where ω represents the final angular velocity of the sphere. Observe that if we take the point of contact between the sphere and the table to be our origin, the angular momentum of the system is conserved as all forces (most notably impulsive forces) pass through this point. The initial angular momentum of the system about this origin is $I\omega_0$. After the collision, the angular momentum is $mru_x + I\omega$. Hence,

$$I\omega_0 = mru_x + I\omega.$$

Applying the non-slip condition $u_x = r\omega$,

$$u_x = \frac{2}{7} r\omega_0.$$

Thus,

$$\theta = \tan^{-1} \frac{u_x}{v} = \tan^{-1} \frac{2r\omega_0}{7\sqrt{2gh}}.$$

21. Bouncing Mass**

Let v_i and u_i be the vertical and horizontal speeds of the mass immediately after the ith collision. Then,

$$\begin{aligned} v_i &= ev_{i-1} \\ \implies v_i &= e^i v_0. \end{aligned}$$

Thus, the impulse delivered by the normal force on the ball due to the ground during the ith collision is

$$I_N = \int N dt = m(v_i + v_{i-1}) = (1+e)mv_{i-1}.$$

Then, the maximum impulse (as the mass may stop moving before the maximum impulse is completely imparted) delivered by the friction force on the mass is then

$$I_f = \int f dt = -\mu \int N dt = -\mu(1+e)mv_{i-1}.$$

Therefore, the minimum final horizontal velocity after the ith collision can be determined by the impulse-momentum theorem.

$$mu_i \geq mu_{i-1} + I_f,$$

where the minimum value is always taken if it is positive. After n collisions, u_n which is given by

$$\begin{aligned} u_n &\geq u_{n-1} - \mu(1+e)v_{n-1} \\ &= u_{n-2} - \mu(1+e)v_{n-1} - \mu(1+e)v_{n-2} \\ &= u_{n-3} - \mu(1+e)v_{n-1} - \mu(1+e)v_{n-2} - \mu(1+e)v_{n-3} \end{aligned}$$

$$= u_0 - \mu(1+e) \cdot \sum_{k=0}^{n-1} v_k$$

$$= u_0 - \mu(1+e) \cdot \sum_{k=0}^{n-1} e^k v_0$$

$$= u_0 - \mu(1+e) v_0 \frac{1-e^n}{1-e}$$

becomes 0. Thus, n is the minimum positive integer for which

$$u_0 - \mu(1+e) v_0 \frac{1-e^n}{1-e} \leq 0$$
$$\implies n \geq \log_e \left| 1 - \frac{u_0(1-e)}{\mu(1+e)v_0} \right|,$$

where e refers to the coefficient of restitution and not Euler's constant. Note that $e < 1$ which causes the inequality sign to reverse when taking $\log_e$ on both sides of a preceding inequality. The minimum positive integer is thus

$$n = \left\lceil \log_e \left| 1 - \frac{u_0(1-e)}{\mu(1+e)v_0} \right| \right\rceil .$$

22. Three Masses**

Define the x and y axes to be positive rightwards and upwards, respectively, and the z-axis to be positive in the direction pointing out of the page such that anti-clockwise rotations are positive. Number the masses m, $2m$ and $3m$ from 1 to 3 in ascending order. Let the final velocity of m be

$$v_1 = \begin{pmatrix} 0 \\ v_y \\ 0 \end{pmatrix} .$$

Then, let ω_1 and ω_2 be the angular velocities of $2m$ and $3m$ with respect to m, immediately after the collision. Choosing m as the reference point for

$\boldsymbol{v} = \boldsymbol{v}_{ref} + \boldsymbol{\omega} \times \boldsymbol{r}$, the final velocities of $2m$ and $3m$ are

$$v_2 = \begin{pmatrix} 0 \\ v_y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \omega_1 \end{pmatrix} \times \begin{pmatrix} -l\cos\alpha \\ l\sin\alpha \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -\omega_1 l \sin\alpha \\ v_y - \omega_1 l \cos\alpha \\ 0 \end{pmatrix},$$

$$v_3 = \begin{pmatrix} 0 \\ v_y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \omega_2 \end{pmatrix} \times \begin{pmatrix} -l\sin\alpha \\ -l\cos\alpha \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} \omega_2 l \cos\alpha \\ v_y - \omega_2 l \sin\alpha \\ 0 \end{pmatrix}.$$

Firstly, the total momentum of the three masses in the y-direction must still be zero after the collision as there is no friction between the wall and m. Then,

$$6v_y = 2\omega_1 l \cos\alpha + 3\omega_2 l \sin\alpha.$$

Now, we require two more equations which can be obtained from the fact that the momenta of masses $2m$ and $3m$ must be conserved in the directions perpendicular to the corresponding rods due to the absence of tensions in the transverse directions. Then,

$$v_y \cos\alpha - \omega_1 l = u \sin\alpha,$$
$$\omega_2 l - v_y \sin\alpha = u\cos\alpha.$$

Solving these three equations simultaneously would yield

$$v_y = \frac{u\sin\alpha\cos\alpha}{3+\cos^2\alpha},$$
$$\omega_1 l = -\frac{3u\sin\alpha}{3+\cos^2\alpha},$$
$$\omega_2 l = \frac{4u\cos\alpha}{3+\cos^2\alpha}.$$

The total impulse delivered by the wall can be computed via the change in the total horizontal momentum of the system of masses. The final total horizontal momentum is

$$p'_x = mv_{1x} + 2mv_{2x} + 3mv_{3x} = \frac{6 + 6\cos^2\alpha}{3 + \cos^2\alpha}mu.$$

Thus, the required impulse J is

$$J = \Delta p_x = \frac{6 + 6\cos^2\alpha}{3 + \cos^2\alpha}mu - 6mu = -\frac{12}{3 + \cos^2\alpha}mu.$$

23. N Disks***

Let the angular velocities ω_i be defined to be positive in the clockwise direction. Let M_i be the total anti-clockwise angular impulse exerted on the ith disk by the $(i+1)$th disk. Then, fM_i would be the total anti-clockwise angular impulse exerted on the $(i+1)$th disk by the ith disk — the factor f stems from the change in radius. Next, let the final clockwise angular velocity of the first disk be Ω. For the system to equilibrate, the disks must not slip relative to each other. Therefore, the final angular velocity of the ith disk must be $(-1)^{i-1}\frac{\Omega}{f^{i-1}}$. If we compare the final and initial velocities of each disk, we obtain N equations.

$$I_1(\omega_1 - \Omega) = M_1,$$

$$I_2\left(\omega_2 + \frac{\Omega}{f}\right) = fM_1 + M_2,$$

$$I_3\left(\omega_3 - \frac{\Omega}{f^2}\right) = fM_2 + M_3,$$

and so on with the ith equation, where $1 < i < N$, being

$$I_i\left(\omega_i + (-1)^i\frac{\Omega}{f^{i-1}}\right) = fM_{i-1} + M_i,$$

and the Nth equation being

$$I_N\left(\omega_N + (-1)^N\frac{\Omega}{f^{N-1}}\right) = fM_{N-1}.$$

Observe that the Nth equation minus f times of the $(N-1)$th equation plus f^2 times of the $(N-2)$th equation minus f^3 times of the $(N-3)$th equation and so on produces a null result on the right-hand side. Then,

$$\sum_{i=1}^{N}(-1)^{i+1}I_i \cdot f^{N-i}\left(\omega_i + (-1)^i\frac{\Omega}{f^{i-1}}\right) = 0.$$

This is valid for both odd and even values of N as their corresponding equations only differ by a minus sign. Now, divide the entire equation by I_1.

Since $\frac{I_i}{I_1} = f^{2i-2}$,

$$\sum_{i=1}^{N}(-1)^{i+1}f^{N+i-2}\left(\omega_i + (-1)^i\frac{\Omega}{f^{i-1}}\right) = 0.$$

Dividing the summation into two parts,

$$\begin{aligned}\sum_{i=1}^{N}(-1)^{i+1}f^{N+i-2}\omega_i &= \sum_{i=1}^{N}f^{N-1}\Omega\\ Nf^{N-1}\Omega &= \sum_{i=1}^{N}(-1)^{i+1}f^{N+i-2}\omega_i\\ \Omega &= \frac{\sum_{i=1}^{N}(-1)^{i+1}f^{i-1}\omega_i}{N}.\end{aligned}$$

24. Rebounding Mass***

Let v_n denote the velocity of the mass immediately after the nth collision, v_n' denote the velocity of the mass immediately before the $(n+1)$th collision and x_n be the distance traveled by the mass up the ramp, until it attains its peak, between its nth and $(n+1)$th collisions. Let the friction force on the mass be $f = \mu mg\cos\theta$. Then, we can obtain a relationship between x_n and v_n based on the work-energy theorem.

$$\begin{aligned}\frac{1}{2}mv_n^2 &= mg\sin\theta x_n + fx_n\\ x_n &= \frac{mv_n^2}{2(mg\sin\theta + f)} = \frac{v_n^2}{2g(\sin\theta + \mu\cos\theta)}.\end{aligned}$$

We can also relate v_n and v_n' in a similar fashion

$$\frac{1}{2}mv_n^2 - 2fx_n = \frac{1}{2}mv_n'^2.$$

From the previous two equations, we obtain

$$v_n'^2 = kv_n^2,$$

where

$$k = \left(\frac{mg\sin\theta - f}{mg\sin\theta + f}\right) = \frac{\sin\theta - \mu\cos\theta}{\sin\theta + \mu\cos\theta}.$$

Furthermore, based on the definition of the coefficient of restitution, we have

$$\begin{aligned} v_{n+1} &= ev_n' \\ \implies v_{n+1}^2 &= ke^2 v_n^2, \\ x_{n+1} &= ke^2 x_n. \end{aligned}$$

Now we just need our base case for x which is given by $x_0 = l$ (you can imagine the initial state of the system as a particle that attains its peak between its 0^{th} and 1^{st} collision). The total distance traveled by the mass is then

$$\begin{aligned} \Delta x &= x_0 + 2x_1 + 2x_2 + 2x_3 + \cdots \\ &= l + 2ke^2 l + 2k^2 e^4 l + \cdots \\ &= 2l(1 + ke^2 + k^2 e^4 + \cdots) - l \\ &= \frac{2l}{1 - ke^2} - l \\ &= \frac{1 + ke^2}{1 - ke^2} l, \end{aligned}$$

where $k = \frac{\sin\theta - \mu\cos\theta}{\sin\theta + \mu\cos\theta}$.

25. Drag Force on Sheet*

Let the total area of the sheet be A. In time dt, the sheet collides with $dm = \rho A v dt$ amount of mass which gains $2vdm$ amount of momentum as they leave the sheet at $2v$ (the change in velocity can be computed in the sheet's frame where the particles approach at $-v$ and rebound at v). The change in momentum of the sheet during this time interval is negative of this. Thus,

$$\begin{aligned} dp &= -2\rho A v^2 dt \\ F_{drag} &= \frac{dp}{dt} = -2\rho A v^2 \\ \frac{F_{drag}}{A} &= -2\rho v^2. \end{aligned}$$

26. Propeling a Car**

Let the instantaneous velocity of the car at a point in time be $v(t)$. Now imagine that you throw a mass dm of baseballs at the car. In the frame which is traveling at v, the baseballs approach the car at a velocity $(u - v)$

and rebound with velocity $(v - u)$. The change in the momentum of the snowball $((2v - 2u)dm)$ is the negative of the change in the momentum of the car in the moving frame. Thus, the change in the momentum of the car in the moving frame is

$$dp = 2(u - v)dm$$
$$F' = \frac{dp}{dt} = 2(u - v)\frac{dm}{dt} = F,$$

where F' is the force on the car in the moving frame and F is the force on the car in the lab frame. They are equal as a consequence of the properties of Galilean transformations. Furthermore, note that $\frac{dm}{dt} \neq \sigma$ where $\frac{dm}{dt}$ is the mass rate of baseballs colliding with the car. This is because, though the baseballs leave your hand at velocity u, they travel at $(u - v)$ relative to the cart. Think of this as something analogous to the Doppler effect. Quantitatively, if adjacent snowballs were separated by a distance l, the time between collisions would be $\frac{l}{u-v}$ as compared to $\frac{l}{u}$ if the cart were stationary. Thus,

$$\frac{dm}{dt} = \frac{u - v}{u}\sigma,$$
$$M\frac{dv}{dt} = 2\frac{(u - v)^2}{u}\sigma,$$
$$\int_0^v \frac{1}{(u - v)^2}dv = \frac{2\sigma}{Mu}\int_0^t dt$$
$$\frac{1}{u - v} - \frac{1}{u} = \frac{2\sigma t}{Mu}$$
$$v = u - \frac{u}{1 + \frac{2\sigma t}{M}},$$
$$\int_0^x dx = \int_0^t u - \frac{u}{1 + \frac{2\sigma t}{M}}dt$$
$$x = ut - \frac{Mu}{2\sigma}\ln\left(1 + \frac{2\sigma t}{M}\right).$$

27. Drag Force on Sphere**

Let the sphere travel at a velocity v in the positive y-direction. In the sphere's frame, the incoming particles travel at a velocity $-v$. Consider the collision between a particle and an infinitesimal surface element on the sphere at azimuthal coordinate ϕ and an angle θ from the z-axis in spherical coordinates. The particle will experience a change in velocity equal to twice of the negated component of its initial velocity along the area vector of the surface element (which is radially outwards). The unit area vector is

$$\hat{\boldsymbol{r}} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}.$$

Thus, the component of the sand's initial velocity along this is

$$\begin{pmatrix} 0 \\ -v \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} = -v\sin\theta\sin\phi.$$

The change in the particle's velocity is -2 times in magnitude of this, in the radial direction.

$$\Delta\boldsymbol{v} = 2v\sin\theta\sin\phi\hat{\boldsymbol{r}}.$$

The change in momentum of the particle, of mass dm, is then

$$dm\Delta\boldsymbol{v} = 2v\sin\theta\sin\phi\hat{\boldsymbol{r}} \cdot dm.$$

The change in momentum of the sphere is negative of this by the conservation of momentum.

$$d\boldsymbol{p} = -2v\sin\theta\sin\phi\hat{\boldsymbol{r}} \cdot dm.$$

Now, we simply have to determine how much mass collides with the infinitesimal surface element at ϕ, θ in a time interval dt. The volume swept by this surface element (in the lab frame now) in time dt is the dot product of the velocity of the sphere and the area vector of the surface element, multiplied by dt. Thus,

$$\begin{aligned} dm &= \rho\boldsymbol{v} \cdot dA\hat{\boldsymbol{r}}dt \\ &= \rho \begin{pmatrix} 0 \\ v \\ 0 \end{pmatrix} \cdot \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} r^2\sin\theta d\phi d\theta dt \\ &= \rho v r^2\sin^2\theta\sin\phi d\phi d\theta dt. \end{aligned}$$

Hence,

$$\frac{d\boldsymbol{p}}{dt} = -2\rho v^2 r^2\sin^3\theta\sin^2\phi\hat{\boldsymbol{r}}d\phi d\theta.$$

The total drag force on the sphere will only be along the y-direction by symmetry. Thus, we can simply compute the y-component of the above.

$$\frac{dp_y}{dt} = -2\rho v^2 r^2 \sin^4\theta \sin^3\phi d\phi d\theta.$$

The total drag force on the sphere is obtained by integrating the above over half the surface of the sphere (as only half collides with the particles).

$$F_{drag} = -\int_0^\pi \int_0^\pi 2\rho v^2 r^2 \sin^4\theta \sin^3\phi d\phi d\theta = -\pi\rho v^2 r^2.$$

Interestingly, comparing this result with the drag force of a sheet shows that the "effective area" of a sphere turns out to be half its cross-sectional area, $\frac{1}{2}\pi r^2$.

28. Carrying Sand**

At steady state, the kinetic energy of the circle plus the sand stuck to it is constant. There are two factors which affect the kinetic energy of this combined system. Firstly, the falling sand collides with the circle and imparts in it kinetic energy. Secondly, the torque about the center due to the weight of the sand clinging on the circle engenders an angular acceleration — a more enlightening perspective of this is to observe that the decrease in gravitational potential energy of the sand is equal to the increase in kinetic energy of the combined system. Let us compute the changes in kinetic energy due to these factors separately and then add them together. For the first factor, though both energy and linear momentum are not conserved during the inelastic collision (the latter is because of the impulsive force due to the axle), angular momentum is conserved about the axle (where the impulsive force due to the axle generates no angular impulse). Let the instantaneous angular momentum and moment of inertia of the combined system about the center of the circle at steady state be L and I respectively. I can be computed from the fact that the linear mass density of the sand on the surface of the circle at steady state should be $\lambda = \frac{\sigma}{r\omega}$ as σdt mass of sand drops on to an arc of length $r\omega dt$ in time dt. This implies

$$I = \pi r\lambda \cdot r^2 = \frac{\pi\sigma r^2}{\omega}.$$

The speed of incoming sand is $v = \sqrt{2gh}$. By the conservation of angular momentum, the angular momentum of the combined system at a time dt

from the current instance is

$$L + dL = L + \sigma r \sin\theta v dt \implies dL = \sigma r \sin\theta \sqrt{2gh} dt,$$

where the change stems from the angular momentum of the falling sand. The new instantaneous moment of inertia $I + dI$ is

$$I + dI = I + r^2 dm \implies dI = \sigma r^2 dt.$$

Since the total kinetic energy of the combined system is $T = \frac{L^2}{2I}$, the change in kinetic energy due to the collision event is

$$\begin{aligned} dT_1 &= \frac{(L+dL)^2}{2I\left(1+\frac{dI}{I}\right)} - \frac{L^2}{2I} \\ &\approx \frac{L^2 + 2LdL}{2I}\left(1 - \frac{dI}{I}\right) - \frac{L^2}{2I} \\ &= \frac{L}{I}dL - \frac{L^2 dI}{2I^2} \\ &= \omega dL - \frac{1}{2}\omega^2 dI \\ &= \left(\sigma r \sin\theta \sqrt{2gh}\omega - \frac{1}{2}\sigma r^2 \omega^2\right) dt. \end{aligned}$$

Proceeding with the second factor, the circle rotates an angle $d\theta = \omega dt$ in time dt at steady state. Then, $\lambda r \omega dt$ amount of sand is effectively transferred from the point of collision to the point where sand is ejected — implying a loss in gravitational potential energy of $2\lambda r^2 g \cos\theta \omega dt$. By the conservation of energy, the increase in the kinetic energy of the combined system due to the second factor is correspondingly

$$dT_2 = 2\lambda r^2 g \cos\theta \omega dt = 2\sigma r g \cos\theta dt.$$

At steady state, the total change in kinetic energy is zero.

$$\begin{aligned} \implies \sigma r \sin\theta \sqrt{2gh}\omega - \frac{1}{2}\sigma r^2 \omega^2 + 2\sigma r g \cos\theta = 0 \\ \omega^2 - \frac{2\sqrt{2gh}\sin\theta}{r}\omega - \frac{4g\cos\theta}{r} = 0, \end{aligned}$$

Solving,

$$\omega = \frac{\sqrt{2gh}\sin\theta}{r} + \sqrt{\frac{2gh\sin^2\theta}{r^2} + \frac{4g\cos\theta}{r}},$$

where the negative solution has been rejected.

29. Sweeping Duster***

The total mass of the duster plus collected-dust system is $m = M + \sigma l x$. The net force on the duster plus dust system at an instant is $mg\sin\theta$. Applying the impulse-momentum theorem across a time interval dt,

$$mg\sin\theta dt = (m + dm)(v + dv) - mv$$

$$\implies mg\sin\theta = \frac{dm}{dt}v + m\frac{dv}{dt}.$$

Observe that

$$\frac{dm}{dt}v + m\frac{dv}{dt} = \frac{dm}{dx}\cdot\frac{dx}{dt}v + m\dot{v}$$
$$= \sigma l v^2 + m\dot{v}.$$

Hence

$$(M + \sigma l x)g\sin\theta = \sigma l v^2 + (M + \sigma l x)\dot{v}.$$

If we let $y = \frac{M}{\sigma l} + x$,

$$yg\sin\theta = \dot{y}^2 + y\ddot{y}.$$

Using $\ddot{y} = \frac{1}{2}\frac{d\dot{y}^2}{dy}$,

$$\dot{y}^2 + \frac{1}{2}y\frac{d\dot{y}^2}{dy} = yg\sin\theta.$$

Multiplying by the integrating factor $2y$,

$$2y\dot{y}^2 + y^2\frac{d\dot{y}^2}{dy} = \frac{d(y^2\dot{y}^2)}{dy} = 2y^2 g\sin\theta$$

$$\int_0^{y^2\dot{y}^2} d(y^2\dot{y}^2) = \int_{\frac{M}{\sigma l}}^{y} 2y^2 g\sin\theta dy$$

$$y^2\dot{y}^2 = \frac{2}{3}y^3 g\sin\theta - \frac{2M^3}{3\sigma^3 l^3}g\sin\theta$$

$$\dot{y} = \sqrt{\frac{2}{3}yg\sin\theta - \frac{2M^3 g\sin\theta}{3\sigma^3 l^3 y^2}},$$

$$\dot{x} = \sqrt{\frac{2}{3}\left(\frac{M}{\sigma l} + x\right)g\sin\theta - \frac{2M^3 g\sin\theta}{3\sigma^3 l^3(\frac{M}{\sigma l} + x)^2}}.$$

30. Raindrop***

Define ρ and λ as the uniform mass densities of the raindrop and the water droplets in the cloud, respectively and let $r(t)$, $m(t)$ and $v(t)$ be the instantaneous radius, mass and velocity of the raindrop. We need three equations to solve for these 3 variables and these comprise 2 equations for $\dot{m}$ and the impulse-momentum theorem. Firstly, the only force acting on the raindrop at an instant in time is the gravitational force. Taking downwards to be positive,

$$\begin{aligned} mgdt &= (m+dm)(v+dv) - mv \\ mg &= \dot{m}v + m\dot{v}. \end{aligned}$$

Furthermore, following from the assumption that our raindrop constantly takes the form of a sphere,

$$m = \frac{4}{3}\pi r^3 \rho \implies \dot{m} = 4\pi r^2 \dot{r} \rho.$$

We can derive another expression for $\dot{m}$ by observing that in a time interval dt, the raindrop sweeps a volume equal to its cross-sectional area multiplied by vdt. Thus, the gain in mass dm in an interval dt is

$$dm = \lambda dV = \lambda \pi r^2 vdt \implies \dot{m} = \lambda \pi r^2 v.$$

Comparing the two expressions for $\dot{m}$,

$$\begin{aligned} v &= \frac{4\rho}{\lambda}\dot{r}, \\ \dot{v} &= \frac{4\rho}{\lambda}\ddot{r}. \end{aligned}$$

Substituting the appropriate expressions into the equation obtained from the impulse-momentum theorem,

$$\begin{aligned} \frac{4}{3}\pi r^3 \rho g &= 4\pi r^2 \dot{r}\rho \cdot \frac{4\rho}{\lambda}\dot{r} + \frac{4}{3}\pi r^3 \rho \cdot \frac{4\rho}{\lambda}\ddot{r} \\ \frac{g\lambda}{4\rho} r &= 3\dot{r}^2 + r\ddot{r}. \end{aligned}$$

Using $\ddot{r} = \frac{1}{2}\frac{d\dot{r}^2}{dr}$,

$$3\dot{r}^2 + \frac{1}{2} r \frac{d\dot{r}^2}{dr} = \frac{g\lambda}{4\rho} r.$$

Multiplying the above by the integrating factor $2r^5$,

$$6r^5\dot{r}^2 + r^6\frac{d\dot{r}^2}{dr} = \frac{g\lambda}{2\rho}r^6$$

$$\int_0^{r^6\dot{r}^2} d(r^6\dot{r}^2) = \int_0^r \frac{g\lambda}{2\rho}r^6 dr \implies r^6\dot{r}^2 = \frac{g\lambda}{14\rho}r^7$$

$$\dot{r} = \sqrt{\frac{g\lambda}{14\rho}r} \implies \int_0^r \frac{1}{\sqrt{r}}dr = \int_0^t \sqrt{\frac{g\lambda}{14\rho}}dt$$

$$2\sqrt{r} = \sqrt{\frac{g\lambda}{14\rho}}t \implies r = \frac{g\lambda}{56\rho}t^2.$$

This gives the acceleration of the raindrop, $\dot{v}$, as

$$\dot{v} = \frac{4\rho}{\lambda}\ddot{r} = \frac{g}{7},$$

which is independent of ρ, λ and t. A faster but less rigorous way of solving

$$mg = \dot{m}v + m\dot{v}$$

is to deduce from the problem statement that $\dot{v}$ should be a constant a (such that $v = at$) since the question did not ask for the acceleration as a function of time. Since $v = \frac{4\rho}{\lambda}\dot{r}$, $r \propto t^2$ such that $m = \frac{4}{3}\pi r^3\rho$ implies that $m = kt^6$ for some constant k. Substituting $m = kt^6$, $\dot{v} = a$ and $v = at$,

$$kt^6 g = 6kt^5 \cdot at + kt^6 a = 7kt^6 a$$

$$a = \frac{g}{7}.$$

This page intentionally left blank

Chapter 7

Statics

As its nomenclature implies, statics refers to the study of objects that are immobile, both translationally and rotationally. Such immobile objects are known to be under a state of static equilibrium. The conditions for static equilibrium are directly linked to the concepts of forces and torques which have been covered in the past few chapters. In a similar vein, objects moving at a constant translational velocity and rotating at a constant angular velocity are said to be under dynamic equilibrium. However, our analysis will mostly be restricted to stationary set-ups as systems with uniformly translating and rotating components are often not meaningful — most of the time, the forces and torques in the system change as the system evolves, such that the dynamic equilibrium is often transient.

7.1 Equilibrium

Translational Equilibrium

For a system to not accelerate translationally, it must not be acted upon by a net external force.

$$\sum \boldsymbol{F} = 0. \tag{7.1}$$

Consider a tug-of-war between two people (Fig. 7.1).

If the forces exerted on the massive rope due to the two people are balanced (i.e. $F_1 = F_2$), the rope does not accelerate. Otherwise, the person who exerts a greater force will win the game as the rope will accelerate towards him or her. Of course, sufficient static friction must exist to "glue" the two to the ground too.

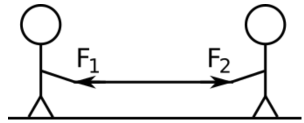


Figure 7.1: Tug-of-war

Rotational Equilibrium

An extended body that is not undergoing a translational acceleration may still not be "static" due to its possible rotational motion.

If the two people then decide to play their game of tug-of-war with a pole and exert equal forces on both sides, the initially stationary pole will definitely undergo angular acceleration and begin to rotate even though its center of mass does not move.

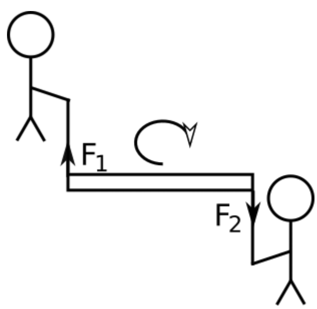


Figure 7.2: Tug-of-war 2

For a system to be in rotational equilibrium, there must be no net torque due to external forces on the system about **every point**.[1]

$$\sum \boldsymbol{\tau} = 0. \tag{7.2}$$

Note that internal forces, which obey the strong law of action and reaction — that requires the lines of actions of an action and reaction pair to be coincident — produce no net torque in a system.

Thankfully, when an object is in a state of translational equilibrium, there is no need to show that the net torque about every point is zero. It is sufficient to show that the net torque about **a single point**, which can be

[1]This is because the angular momentum of a static system is required to be constant about any point. A change in the angular momentum of a system about any point definitely implies that a particle has undergone acceleration.

located anywhere, is zero. This is due to the fact that the net torques about any two points are the same if the vector sum of forces is zero.

Proof: Consider two origins O and O$'$. $\boldsymbol{R}$ is the vector pointing from O$'$ to O. Let $\boldsymbol{r}_i$ and $\boldsymbol{r}'_i$ be the vectors from origins O and O' to the point of action of the ith force $\boldsymbol{F}_i$ respectively (Fig. 7.3).

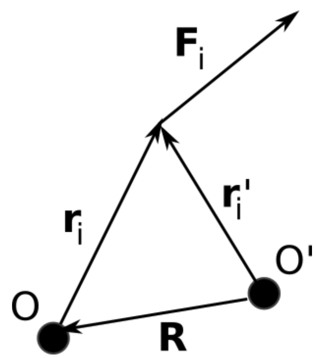


Figure 7.3: Torques about two origins O and O'

Then, if there are a total of N forces on the system, the net torque about origin O is

$$\sum \boldsymbol{\tau} = \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i.$$

The net torque about origin O' is

$$\begin{aligned}
\sum \boldsymbol{\tau}' &= \sum_{i=1}^{N} \boldsymbol{r}'_i \times \boldsymbol{F}_i \\
&= \sum_{i=1}^{N} (\boldsymbol{r}_i + \boldsymbol{R}) \times \boldsymbol{F}_i \\
&= \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i + \boldsymbol{R} \times \left(\sum_{i=1}^{N} \boldsymbol{F}_i \right) \\
&= \sum_{i=1}^{N} \boldsymbol{r}_i \times \boldsymbol{F}_i \\
&= \sum \boldsymbol{\tau}
\end{aligned}$$

since $\sum_{i=1}^{N} \boldsymbol{F}_i = 0$. Hence, the net torques about any two points of a system in translational equilibrium are equal. In light of this property, we should always wisely pick our origin at locations that exclude as many torques, which arise due to unknown forces, as possible, to swiftly solve for our desired variables. This is in stark contrast with the stance in rotational dynamics where the origin is predominantly chosen at the center of mass or an ICoR.

Static Equilibrium

For a system to be in a state of static equilibrium, it must be under both translational and rotational equilibrium and thus must satisfy both of their respective conditions. Such static situations form the core of this chapter. The crux of analyzing static set-ups involves considering appropriate systems and sub-systems and balancing the external forces and torques on them. The importance of free-body diagrams in this process cannot be understated.

General Procedure for Solving Static Systems

The general procedure in analyzing an n-dimensional static system is to write down n equations ($\sum F_q = 0$ where q indicates a certain direction) that balance the components of forces along n independent directions which are not necessarily perpendicular and write down a number of torque balance equations ($\sum \tau_q = 0$) about judiciously-chosen origins corresponding to the possible number of rotational motions of the system. In the common case where $n = 2$, we need 2 force balance equations and 1 torque balance equation about a certain origin. However, any number of force balance equations can be replaced by an equivalent number of torque balance equations about different origins. For example, we can have 1 force balance equation and 2 torque balance equations about different origins or 3 torque balance equations about 3 different origins that are not collinear. To show this, suppose we write $\sum \boldsymbol{\tau} = \mathbf{0}$ and $\sum \boldsymbol{\tau}' = \mathbf{0}$ with respect to two different origins O and O' (see diagram in previous section). Note that we use vector notation instead of taking the component in a particular direction as there can only be a single rotational direction in a two-dimensional problem. Subtracting the former equation from the latter,

$$\sum \boldsymbol{\tau}' - \sum \boldsymbol{\tau} = \boldsymbol{R} \times \left(\sum_{i=1}^{N} \boldsymbol{F}_i \right) = \mathbf{0},$$

where $\boldsymbol{R}$ is the position vector of O with respect to O'. The above equation is equivalent to stating that the sum of forces perpendicular to $\boldsymbol{R}$

is zero — analogous to a force balance equation. Therefore, an additional torque balance equation about an origin whose position vector relative to other already-chosen origins is not parallel to other relative vectors between already-chosen origins generates at least one additional force balance equation along a new direction — implying that force balance equations can be superseded by torque balance equations. However, note that there must always be a torque balance equation as the total number of linearly independent directions (for force balance) is limited.

Trading a force balance equation for a torque balance equation is actually very favourable. The reason is that by balancing the sum of forces along a particular direction, we can usually only eliminate a single force (by choosing the direction of interest to be perpendicular to this force). However, by taking torques about an origin that is the intersection of various forces, many forces can be eliminated in one fell swoop. We shall see how we can exploit this fact later, but let us first practise balancing forces and torques in a few simple systems.

Systems as a Whole

Problem: A rope is attached to a vertical pole. Then, a horizontal force F is exerted on the left end of the rope. The rope remains stationary and its right end makes an angle of $30°$ with the vertical pole. Determine the mass of the rope, m.

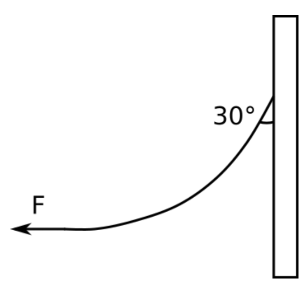


Figure 7.4: Rope attached to pole

This simple example serves to highlight an important fact about ropes. A rope segment cannot withstand or exert a force perpendicular to its instantaneous gradient as it will deform under such a force (or its reaction) — implying that the tension in an infinitesimal segment of a immobile rope can only exist in the longitudinal direction. Therefore, we can conclude that the force on the rope at its right end subtends $30°$ with the vertical. Drawing a free-body diagram of the rope, we obtain

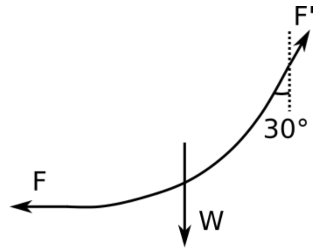


Figure 7.5: Free-body diagram of rope

where F' is the force on the rope due to the pole. Note that the center of mass of the rope may not be at the midpoint of the rope. It does not even need to be on the rope. Balancing forces in the horizontal and vertical directions respectively,

$$F = F' \sin 30^\circ$$
$$W = F' \cos 30^\circ = mg$$
$$\implies m = \frac{\sqrt{3}F}{g}.$$

Problem: There is a man of mass m who walks on a uniform pole of mass M and length l. The right end of the pole is attached to the ceiling via a massless rope that makes an angle of 60° with the pole. On the other hand, there is a hinge H at the left end. If the rope snaps when the tension reaches a critical limit T_0, what is the maximum distance from the left end of the pole that the man can attain if the pole remains static throughout the whole process?

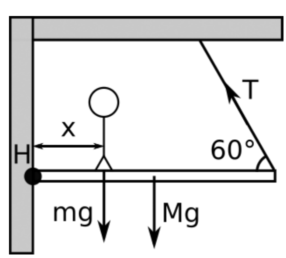


Figure 7.6: Man on pole

Let the current position of the man be a distance x from the left end of the pole. Balancing torques on the pole-and-man system about the hinge H,

$$T \sin 60^\circ \cdot l = mgx + Mg\frac{l}{2}.$$

Thus, when the tension force reaches its maximum value T_0, the maximum distance that the man travels x_{max} is given by

$$x_{max} = \frac{(\sqrt{3}T_0 - Mg)l}{2mg}.$$

Note that we deliberately defined our pivot to be at hinge H as we want to ignore the force on the pole due to the hinge which is not given.

Neat Tricks

When there are three unknown forces, with known directions, in a system and we wish to solve for one of them, the equilibrium equations can be applied in a convenient manner.

Problem: Consider a section of a truss structure in the figure below. All member sections are massless and a force P is applied at the right end of the truss. Determine[2] F_1, F_2 and F_3 if the system remains in static equilibrium.

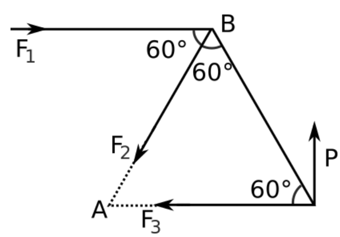


Figure 7.7: Three unknown forces

To solve for F_1, we can extend the lines of action of F_2 and F_3 to their point of intersection A. Then by taking moments about pivot A, the net torque due to F_2 and F_3 is zero. Thus,

$$F_1 = \frac{2P}{\sqrt{3}}.$$

To solve for F_2, notice that F_1 and F_3 are parallel. Then, we can balance forces in the direction perpendicular to them to eliminate the need to consider them.

$$F_2 \cos 30^\circ = P$$
$$F_2 = \frac{2P}{\sqrt{3}}.$$

[2] These forces can only be directed longitudinally along the rods as the rods are two-force members (see section after this).

Having solved for F_1 and F_2, we can solve for F_3 by balancing forces in the horizontal direction. However, the point of this exercise is to determine the forces with the use of a single equation. Hence, we shall consider torques about the point of intersection of F_1 and F_2, denoted by B. Then,

$$F_3 = \frac{P}{\sqrt{3}}.$$

In conclusion, when solving for a force and given two other unknown non-parallel forces, one should take torques about the point of intersection of the two other forces. Next, when solving for a force with two other parallel unknown forces, one should balance forces in the direction perpendicular to those unknown forces.

Forces at Two Points of a System

When there are two net external forces at two points of a system (there may be more than one force at a point but these forces can be vectorially combined into a net force), they must be equal in magnitude and opposite in direction for the system to remain in static equilibrium. Furthermore, the two net forces must be directed along the line joining the two points of action. The first two conditions stem from the fact that the vector sum of the external forces must be zero. The latter results from the constraint that the net torque about any point must be zero. By considering torques about one point of action, it is obvious that the line of action of the net force at the other point of action must pass through this point of action.

An important consequence of this fact pertains to massless members, which are structural components, on which forces act at two different points (depicted on the right of Fig. 7.8). Such members are known as two-force members. As the two points of action are usually the ends of the member, the net forces at the two ends of the member can only be longitudinal — as seen in the previous example.

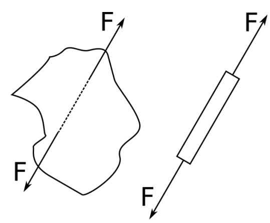


Figure 7.8: Net external force at two different points of a system

Forces at Three Points of a System

If net external forces are exerted with three non-overlapping lines of action on three unique points of a system, the lines of action of the net external forces must be concurrent or parallel for the system to remain static.

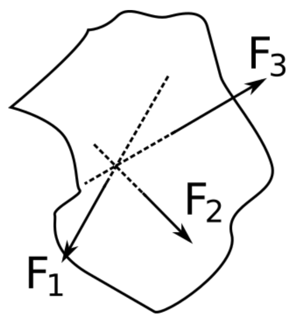


Figure 7.9: Concurrency of forces

This is because, if the three lines of actions are not parallel and are not concurrent, the net torque taken about the point of intersection of two lines of actions will be non-zero due to an external torque created by the other force. Thus, if the lines of actions are non-parallel, they must be concurrent. In the case of three parallel lines of actions, static equilibrium is attainable as the external torques about any point can be balanced. One can also understand this by imagining that these three parallel lines are concurrent at infinity.

This is a useful theorem for determining the possible configurations of a system — acted upon by external forces at three unique points — in which it can remain in static equilibrium.

Problem: A thin uniform pole of length l is placed on two frictionless ramps with angles of inclination α and β respectively. What is/are the possible angle(s) of θ for the pole to remain at static equilibrium?

Observe that there are external forces exerted at three different points on the pole — namely the normal forces at its two ends and a gravitational force at its centroid. It is then necessary for the lines of action of these three forces to intersect in order for the pole to be in static equilibrium.

Define a coordinate system with a vertical y-axis that is positive in the upwards direction and a horizontal x-axis that is positive rightwards. The origin is at O as shown in Fig. 7.10 on the next page. Let the coordinates of the left end of the rod be (x, y) (point A). Then, the coordinates of the right end of the rod (point B) and the centroid are $(x + l\cos\theta, y - l\sin\theta)$ and

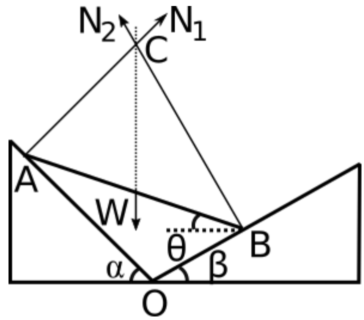


Figure 7.10: Pole on ramps

$(x + \frac{l\cos\theta}{2}, y - \frac{l\sin\theta}{2})$ respectively. In order for the three forces to be concurrent, the x-coordinate of the point of intersection of the two normal forces must be $x + \frac{l\cos\theta}{2}$.

Let us assume that this condition is satisfied and define the coordinates of the point of intersection as $(x+\frac{l\cos\theta}{2}, k)$ (point C). Furthermore, we know that the gradients of the two normal forces N_1 and N_2 are $\frac{1}{\tan\alpha}$ and $-\frac{1}{\tan\beta}$ respectively. Thus, in order for our proposed coordinates of the point of intersection to be correct, the slopes between the points of action of the normal forces and this point of intersection must be commensurate with the aforementioned values.

$$\frac{y - l\sin\theta - k}{x + l\cos\theta - (x + \frac{l\cos\theta}{2})} = -\frac{1}{\tan\beta},$$

$$\frac{k - y}{x + \frac{l\cos\theta}{2} - x} = \frac{1}{\tan\alpha}.$$

Solving the equations above for θ, we obtain

$$\tan\theta = \frac{1}{2}\left(\cot\beta - \cot\alpha\right).$$

There is only one solution for θ in the range $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. For readers who are more geometrically inclined, we can consider the circle that passes through the points A, B and C (Fig. 7.11). Draw a vertical line through C and let its point of intersection with the circle and line AB (the rod) be D and E respectively.

Since angles of the same segment are equal,

$$\angle DBA = \angle DCA = \alpha,$$

$$\angle DAB = \angle DCB = \beta.$$

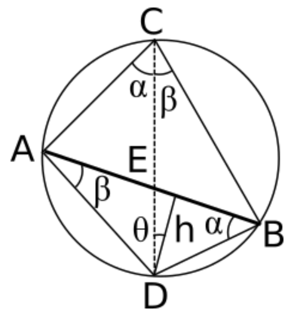


Figure 7.11: Circle passing through points A, B and C

Let h be the perpendicular length from point D onto line AB. For point E to be the midpoint of line AB,

$$h \cot \alpha + h \tan \theta = h \cot \beta - h \tan \theta$$
$$\implies \tan \theta = \frac{1}{2}(\cot \beta - \cot \alpha).$$

Problem: A cuboid lies on a rough, massive inclined plane. If the total normal force on the cuboid can be reduced to an equivalent normal force that only acts at one point on the cuboid, where should that one point be so that the cuboid is able to remain at static equilibrium?

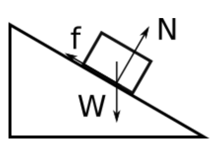


Figure 7.12: Cuboid on inclined plane

The block is again acted upon by three forces, namely the friction due to the plane, its own weight and the normal force due to the plane. Since the friction force is aligned with the slope of the ramp, the normal force on the block due to the plane must act at the point at which a line drawn vertically downwards from the center of mass intersects with the surface of the plane.

7.2 Connected Components

In static set-ups, different components may be related to others in various fashions via diverse connections and supports. The following table in Fig. 7.13 summarizes the common forms of connections and supports.

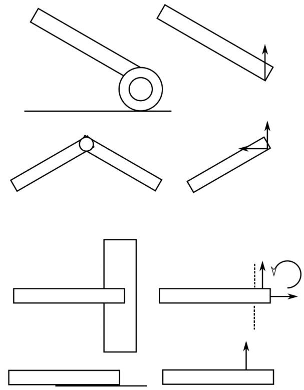


Figure 7.13: Cuboid on inclined plane

The left-hand side of the above figure depicts possible pictorial representations of the corresponding supports and connections while the right illustrates the possible forces and moments on the the members connected to the supports.

- When connected to a roller, a member experiences a force perpendicular to the surface that supports the roller. The member is free to move in the translational direction parallel to the surface and to rotate. A roller is a support and cannot be used to connect structural members.
- In a pinned connection, the connected members can experience a force due to the pin in any direction. The corresponding points of connection on each member cannot move translationally relative to each other (i.e. the members cannot be separated) but the members are able to rotate relative to each other.
- In a fixed connection, the member experiences both forces and moments due to the connection, usually because it is embedded in a medium. The member is neither able to translate nor rotate.
- A simple support is one in which a member leans on a surface. It is similar to the roller support; a force that is perpendicular to the surface acts on the member.

When components in a static system are connected by pinned supports, it becomes necessary to divide the system at the pins and consider the internal forces on the sub-systems due to the pin to generate sufficient equations to entirely solve for the system. This is due to the fact that a pinned connection can only exert forces and not moments. Thus, "dismembering" the system at a pin introduces at most two variables but generates three additional equilibrium equations due to the addition of a new sub-system. If an external force acts on a pin, forces on the pin have to be balanced too so that the forces on the two connected members due to the pin can be related.

Problem: Two sticks of masses m and M are connected by a pin as shown below. They produce a right angle with respect to each other. If a force P is exerted on the pin, determine the reaction forces on the sticks due to the ground. Assume that the ground is sufficiently rough such that the sticks do not slip.

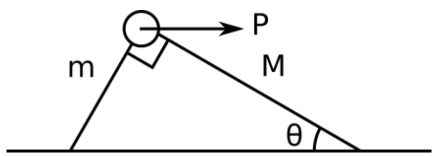


Figure 7.14: Two sticks

We begin by drawing free-body diagrams of the sticks and pin.

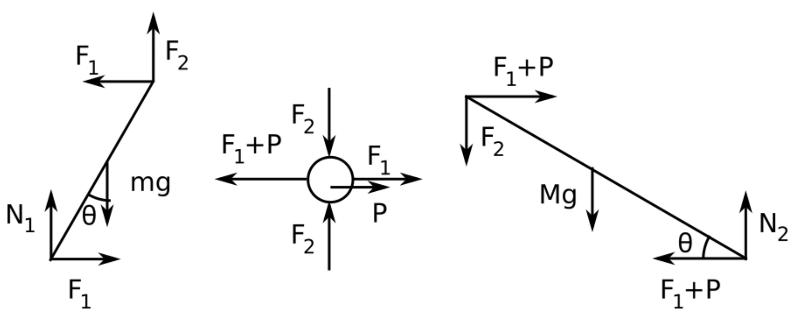


Figure 7.15: Free-body diagrams

If F_1 is the horizontal component of the force on the left stick due to the pin, $F_1 + P$ must be the horizontal component of the force on the right stick due to the pin as the forces on the pin must be balanced. Similarly, if F_2 is the vertical component of the force on the left stick due to the pin, F_2 must also be that on the right stick. We then proceed to label the other forces while accounting for translational equilibrium in the horizontal direction. Balancing forces on the sticks in the vertical direction,

$$N_1 + F_2 = mg,$$
$$N_2 = Mg + F_2.$$

Balancing torques about the bottoms of the sticks,

$$F_1 \cos\theta + F_2 \sin\theta = \frac{mg\sin\theta}{2},$$

$$(F_1 + P)\sin\theta - F_2\cos\theta = \frac{Mg\cos\theta}{2}.$$

Solving,

$$F_1 = \frac{(m+M)g\sin\theta\cos\theta}{2} - P\sin^2\theta,$$

$$F_2 = \frac{mg\sin^2\theta}{2} - \frac{Mg\cos^2\theta}{2} + P\sin\theta\cos\theta,$$

$$N_1 = \frac{mg}{2} + \frac{(m+M)g\cos^2\theta}{2} - P\sin\theta\cos\theta,$$

$$N_2 = \frac{Mg}{2} + \frac{(m+M)g\sin^2\theta}{2} + P\sin\theta\cos\theta,$$

where F_2 is extraneous in a certain sense as it was not asked for in the problem. Notice that if we did not consider the two sticks as two separate systems, there will be four force variables (a horizontal and vertical component for each of the reaction forces due to the ground) but only three equilibrium equations. Hence, we could not have solved for the reaction forces without considering the internal forces of this system.

Now, you may be tempted to think that one should always divide a system into smaller sub-systems to generate more equations. Progressively dividing beams into smaller and smaller sections should produce more and more equations! Well, the obvious trade-off is the increase in the number of variables. In the case of fixed connections, there is an increase in three variables (two components of force and a moment) when dividing a system at a fixed connection. This evens out the increase in equilibrium equations. Furthermore, there is already nothing interesting to solve for as it is already known that the connected members cannot translate or rotate with respect to each other. Hence, dividing a system into smaller sub-systems in this case is not beneficial, but redundant. In the case of a single rigid beam, each section is fixed relative to each other — rendering the consideration of individual sections superfluous. We should only consider sub-systems when there are additional degrees of freedom. A degree of freedom is a possible, independent translational or rotational motion of a component in a system. By independent, we mean that if we keep all other components of the system fixed, the particular component can still move.

7.3 Friction

The previous situations were devoid of static friction. However, in systems containing static friction, there will be a constraint on the relationship between certain physical properties of a system that must be satisfied for the system to remain in static equilibrium. This is due to the upper limit on the magnitude of static friction between two relatively stationary surfaces, $|f| \leq \mu_s N$ where μ_s is the coefficient of static friction and N is the normal force between the surfaces. The absolute value sign is extremely important here as we could have incorrectly guessed the direction of the frictional force which causes f to be negative. Static friction always acts in a direction to oppose impending motion. On a side note, the coefficient of static friction is often greater than that of kinetic friction, $\mu_s > \mu_k$ which explains why heavy objects become easier to push once they begin moving.

If static friction acts on a system, a variable for each frictional force should be defined and solved in terms of the other variables. Only after obtaining the expressions for the frictional forces and the corresponding normal forces should we impose the condition pertaining to the maximum magnitude of friction. Then, a number of inequalities, which is equal to the number of static frictional forces, will be obtained.

Problem: A block of mass m lies on a massive inclined plane with an angle of inclination θ and is attached to another mass M via a frictionless fixed pulley. If the coefficient of static friction between the plane and the block is μ, determine the appropriate conditions that must be satisfied so that the two blocks remain in static equilibrium.

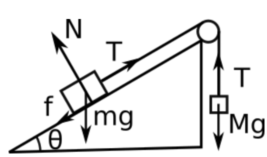


Figure 7.16: Two connected blocks on inclined plane

Firstly, we can identify and label the forces on the two blocks as shown in the diagram above. f refers to the static friction force between the plane and the block on the ramp. The free-body diagrams are superimposed here for convenience. Balancing forces, we obtain

$$T = Mg,$$
$$T = mg \sin\theta + f,$$
$$N = mg \cos\theta.$$

Solving for f,

$$f = Mg - mg\sin\theta.$$

Lastly, we can impose the condition that $|f| \leq \mu N$.

$$|Mg - mg\sin\theta| \leq \mu mg\cos\theta.$$

Now, two separate cases need to be considered. If $M \geq m\sin\theta$,

$$\begin{aligned} Mg - mg\sin\theta &\leq \mu mg\cos\theta \\ M &\leq (\sin\theta + \mu\cos\theta)m. \end{aligned}$$

Thus, the overall condition in this case is

$$m\sin\theta \leq M \leq (\sin\theta + \mu\cos\theta)m.$$

Otherwise if $M \leq m\sin\theta$,

$$\begin{aligned} mg\sin\theta - Mg &\leq \mu mg\cos\theta \\ M &\geq (\sin\theta - \mu\cos\theta)m. \end{aligned}$$

The overall condition in this second case is then

$$(\sin\theta - \mu\cos\theta)m \leq M \leq m\sin\theta.$$

Ultimately, as long as $(\sin\theta - \mu\cos\theta)m \leq M \leq (\sin\theta + \mu\cos\theta)m$, the two masses will remain in static equilibrium.

Problem: A ladder with a uniform mass distribution is placed on a rough ground and leans on a smooth wall. The ladder subtends an angle θ with the horizontal. Determine the conditions required for the ladder to remain in static equilibrium.

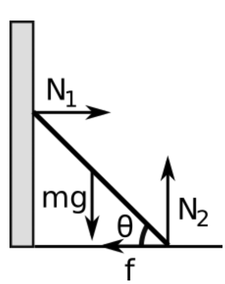


Figure 7.17: Leaning ladder on rough ground

As always, we label the forces on the ladder. Balancing torques about the point of intersection of the two normal forces,

$$fl\sin\theta = mg\cdot\frac{l\cos\theta}{2}$$
$$f = \frac{mg\cot\theta}{2}.$$

Balancing the forces on the ladder in the vertical direction,

$$N_2 = mg.$$

If we impose the condition that $|f| \leq \mu N$,

$$\frac{mg\cot\theta}{2} \leq \mu mg$$
$$\frac{\cot\theta}{2} \leq \mu.$$

We did not need to consider two cases as $f = N_1 > 0$ — we are certain of the direction of friction.

Impending Motion

The previous situations dealt with conditions for static equilibrium in systems with static friction. On the other hand, there are scenarios in which a system is just on the verge of moving or slipping. In such situations, certain static friction forces attain their largest attainable value: $|f| = \mu N$. This leads to a neat condition for surfaces with a single point of contact — the net contact force (comprising friction and the normal force) subtends an angle $\tan^{-1}\mu$ from the normal when the surfaces are about to slip relative to each other. Let us first consider a system with a single friction force.

Problem: A small block of mass m is placed on a massive, rough inclined plane with an angle of inclination θ. If the coefficient of static friction between the block and the plane is $\mu > \tan\theta$ (so that the block does not slip under its own weight), determine the minimum external force required to move the block.

Firstly, we introduce the notion of a force triangle which is applicable when we have three forces and are unconcerned with the torques produced by them. The vector sum of the contact force on the block due to the plane $\boldsymbol{F}$ (normal force-cum-friction), its weight $\boldsymbol{W}$ and the external force $\boldsymbol{P}$ yields the null vector at equilibrium (when the block has yet to slip). Therefore, if the heads and tails of the three pairs of these three vectors are conjoined,

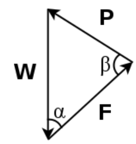


Figure 7.18: Force triangle

a vector triangle must be formed (Fig. 7.18). Its cyclic nature signifies the fact that their sum yields the null vector.

When the block is on the verge of slipping, $\boldsymbol{F}$ is directed at an angle $\tan^{-1}\mu$ from the normal of the slope — implying that the angle α between $\boldsymbol{F}$ and $\boldsymbol{W}$ is either $\theta - \tan^{-1}\mu$ or $\theta + \tan^{-1}\mu$ (depending on the direction of impending motion). Applying the sine rule,

$$\frac{P}{\sin\alpha} = \frac{W}{\sin\beta} \implies P = \frac{W\sin\alpha}{\sin\beta},$$

where β is the angle between $\boldsymbol{F}$ and $\boldsymbol{P}$. It is variable as the magnitudes of $\boldsymbol{F}$ and $\boldsymbol{P}$ are not fixed. Evidently, $|P|$ is minimized when we choose $\alpha = \theta - \tan^{-1}\frac{1}{\mu}$ and $\beta = \frac{\pi}{2}$.

$$\begin{aligned} P_{min} &= mg\sin(\theta - \tan^{-1}\mu) = mg[\sin\theta\cos(\tan^{-1}\mu) - \cos\theta\sin(\tan^{-1}\mu)] \\ &= \frac{mg}{\sqrt{\mu^2+1}}(\sin\theta - \mu\cos\theta). \end{aligned}$$

$\boldsymbol{P}_{min}$ is perpendicular to $\boldsymbol{F}$ at this juncture and subtends an angle $\pi-\alpha-\beta = \frac{\pi}{2} - \theta + \tan^{-1}\mu$ with the vertical. As seen from the fact that this angle is larger than $\frac{\pi}{2}$ or from the fact that the above value of P_{min} is negative, the component of $\boldsymbol{P}_{min}$ parallel to the inclined plane points downslope — agreeing with our intuition that it should be easier to push the block down the slope as gravity is aiding us.

When there are multiple friction forces, the situation becomes more complicated as it is necessary to determine which point slips first or whether different points slip simultaneously. Generally, applying an external force to a system with many frictional forces will not lead to simultaneous slipping.

To determine the minimum external force required to cause a system comprising multiple static frictional forces to slip, one should determine the force required to make each of the points slip. That is, one should plug in $|f| = \mu N$ for the appropriate f and N for each of the individual friction forces, one at a time, and solve for the external force required. Then, the minimum of all computed forces is chosen.

Problem: Referring to the previous problem on Fig. 7.14, if the coefficient of static friction between the ground and the sticks is now μ, determine the minimum P required to make the system slip.

We will use the results that we have previously calculated. Assuming that the left stick slips first,

$$F_1 = \mu N_1,$$
$$P_1 = \frac{g[(m+M)\sin\theta\cos\theta - \mu m - \mu(m+M)\cos^2\theta]}{2(\sin^2\theta - \mu\sin\theta\cos\theta)}.$$

Assuming that the right stick slips first,

$$F_1 + P = \mu N_2,$$
$$P_2 = \frac{g[(m+M)\sin\theta\cos\theta - \mu M - \mu(m+M)\sin^2\theta]}{2(\sin^2\theta + \mu\sin\theta\cos\theta - 1)}.$$

The minimum P is then the smaller of the values of P_1 and P_2.

Next, the assumption of simultaneous slipping is usually made in determining the optimal configuration of systems that are statically indeterminate — a concept that will be explored in a later section. Simply put, there are more force variables than equations. Then, the exact values of these forces can vary according to how the system was constructed. Hence, it is possible to make the points in a statically indeterminate system slip simultaneously — generating additional equations that enable the system to be solvable and leading to situations which are optimal in a certain respect, most of the time. Consider the following example.

Problem: A ladder leans on a rough wall with a coefficient of static friction μ_1 and is placed on a rough ground with a coefficient of static friction μ_2. Determine the minimum angle θ that the ladder makes with the horizontal, for which it does not slip.

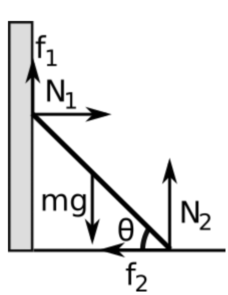


Figure 7.19: Ladder on rough ground and wall

At the minimum angle θ, both friction forces f_1 and f_2 attain their maximum values $\mu_1 N_1$ and $\mu_2 N_2$ respectively. When the ladder slips, it tends to decrease the angle it makes with the horizontal. Hence, the bottom of the ladder tends to slip towards the right while the top of the ladder tends to slip downwards.

$$N_1 = f_2 = \mu_2 N_2,$$
$$f_1 = \mu_1 N_1 = \mu_1 \mu_2 N_2.$$

Balancing forces in the vertical direction,

$$f_1 + N_2 = mg$$
$$\implies N_2 = \frac{mg}{1 + \mu_1 \mu_2}.$$

Taking torques about the bottom of the ladder,

$$N_1 \sin\theta + f_1 \cos\theta = \frac{mg}{2}\cos\theta$$

$$\frac{\mu_2}{1+\mu_1\mu_2} mg \sin\theta + \frac{\mu_1\mu_2}{1+\mu_1\mu_2} mg\cos\theta = \frac{mg}{2}\cos\theta$$

$$\theta = tan^{-1}\left(\frac{\frac{1}{\mu_2} - \mu_1}{2}\right).$$

In most situations, $\mu_1 < 1$ and $\mu_2 < 1$ such that θ is positive. The above is a common method presented by various sources. However, it is not extremely convincing as it is not obvious that the simultaneous attainment of the maximum values of the two frictional forces is possible — though intuition may hint at this. Consider the following proof based on the concurrency of the net forces at the three points of the ladder.

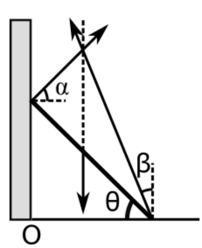


Figure 7.20: Reaction forces and weight

In order for the system to remain static, the lines of action of the two reaction forces at the ends of the ladder and its weight must be concurrent.

Let the origin be at the bottom end of the wall and denote the length of the ladder as l. The coordinate system is such that the upwards and rightwards directions are positive. Then, the coordinates of the top and bottom ends of the ladder are $(0, l\sin\theta)$ and $(l\cos\theta, 0)$ respectively. Since the x-coordinate of the ladder's centroid is $\frac{l\cos\theta}{2}$, let the coordinates of the point of concurrency be $(\frac{l\cos\theta}{2}, y)$. Then, for the three forces to be concurrent at that point,

$$\frac{y - l\sin\theta}{\frac{l\cos\theta}{2}} = \tan\alpha$$
$$\frac{y}{\frac{l\cos\theta}{2}} = \cot\beta.$$

Eliminating y and solving for θ,

$$\tan\theta = \frac{\cot\beta - \tan\alpha}{2}.$$

It can be observed that θ is minimized when $\cot\beta$ is minimized and when $\tan\alpha$ is maximized. Furthermore, from the upper limit on the magnitude of friction at a surface relative to the normal force due to that same surface,

$$\cot\beta \geq \frac{1}{\mu_2},$$
$$\tan\alpha \leq \mu_1.$$

Hence, the minimum value of θ is

$$\theta = \tan^{-1}\left(\frac{\frac{1}{\mu_2} - \mu_1}{2}\right).$$

To see why simultaneous slipping is attainable, consider the case where α has reached its maximum value while β has not. Then, β can be increased further to lower the point of concurrency and hence decrease α — saving the ladder from the verge of slipping. A similar argument can be made in the reverse direction.

Normal Force in Impending Motion

Generally, the normal force between two surfaces with a non-negligible contact area is distributed (not necessarily uniformly) along the entire common surface. Determining the pressure distribution is generally an intractable problem as the force and torque balance equations only yield aggregated values of the force and torques associated with the normal force. However, when a system is on the verge of slipping, the normal force undertakes a boundary case which is often tractable. A frequent condition in such situations is that the entire normal force acts at a single point on the common surface.

Problem: A uniform pencil with a hexagonal cross-section lies on a rough, massive inclined plane with an angle of inclination θ ($\cot\theta \leq \sqrt{3}$). If the coefficient of static friction between the slope and the pencil is $\mu > \frac{1}{\sqrt{3}}$, determine the minimum α for which the pencil can remain at static equilibrium.

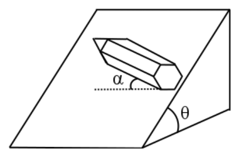


Figure 7.21: Pencil on slope

A common trick in evaluating torques along a certain direction in a three-dimensional problem is to squash the set-up along the direction of the torque (forces which act at points squashed along the same line are vectorially combined) as the extent of the set-up in that direction does not matter in the cross-product $\boldsymbol{r} \times \boldsymbol{F}$ (the component of $\boldsymbol{r}$ parallel to $\boldsymbol{r} \times \boldsymbol{F}$ is inconsequential). Therefore, we can compress the pencil in the direction perpendicular to its cross-section to obtain the following free-body diagram.

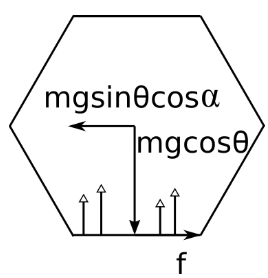


Figure 7.22: Free-body diagram after squashing

Let the mass of the pencil be m. The $mg\cos\theta$ component of the pencil's weight acts at the center of the hexagon in the direction normal to the slope while the $mg\sin\theta\cos\alpha$ component of its weight acts in the plane of the slope and perpendicular to the axis of the pencil. Generally, friction and a normal force also act at the bottom of the pencil's base (the edge at the bottom of the hexagon). Though the normal force is generally distributed along the entire edge (represented by the vectors with white arrowheads), we argue that in the limit where α is minimum, the entire normal force acts at the left end of the bottom edge. To see why this is so, compute torques about the left end of the bottom edge — the anti-clockwise torque due to the leftwards component of the pencil's weight ($mg \sin\theta \cos\alpha$) and the possible anti-clockwise torque due to the normal force on the bottom edge must balance the fixed clockwise torque due to the downwards (normal to the slope) component of the pencil's weight ($mg \sin\theta$). In order for α to be minimized, the former anti-clockwise torque must be maximized while the latter anti-clockwise torque must be minimized — implying that the boundary case occurs when the normal force solely acts at the left end such that it generates zero torque about the left end. At this juncture, the torques due to the two components of the pencil's weight about the left end must be balanced — implying that the vector obtained from combining these two components must pass through the left end. Thus,

$$\tan 30^\circ = \frac{\sin\theta\cos\alpha}{\cos\theta}.$$

The minimum α satisfies

$$\cos\alpha = \frac{\cot\theta}{\sqrt{3}} \implies \alpha = \cos^{-1}\left(\frac{\cot\theta}{\sqrt{3}}\right).$$

The ratio between the friction force f and the normal force N at this juncture is

$$\frac{f}{N} = \frac{mg\sin\theta\cos\alpha}{mg\cos\theta} = \tan\theta\cos\alpha = \frac{1}{\sqrt{3}} < \mu,$$

which shows that the system does not slip.

7.4 Strings under Distributed Force

This section will explore how a string can remain in static equilibrium when acted upon by a force distributed along its entire length. A string "transmits" a force via tension, which means that every section is being pulled by its adjacent sections. A string cannot transmit a compressive force as it will become flimsy when pushed upon and collapses.

The important variables here are the shape of the string and the tension in the string. The common denominator in such problems is the need to consider infinitesimal sections of the string and integrate the relevant variables over the entire string. This is because the tension in a string is an internal force which necessitates the consideration of different segments to find the tension at each point.

Let us begin by considering a simple example. A string of length l and uniform linear mass density λ is hung vertically from a ceiling. The string remains static. Find the tension in the string at a height y from the bottom of the string.

Well, the simple approach would be to consider the bottom segment of the string of length y. The tension on this segment must balance its weight which is λgy. Thus, the tension at that height must be λgy. However, let us make things more interesting by considering infinitesimal segments of the string.

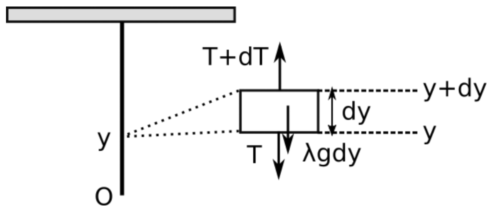


Figure 7.23: Hanging string

Consider an infinitesimal segment of string with ends at vertical coordinates y and $y + dy$ (Fig. 7.23). The y-axis is defined to be positive upwards with the origin O at the bottom of the string. The tension at these ends are T and $T+dT$ respectively. Balancing the forces on this infinitesimal segment,

$$T + dT = T + \lambda gdy$$

$$dT = \lambda gdy.$$

Integrating this,

$$T = \lambda gy + c.$$

Now, we need to impose a boundary condition on the tension to solve for c. We can either use the fact that $T = 0$ at $y = 0$ as there is no mass that needs to be supported at the free bottom end, or $T = \lambda gl$ at $y = l$, as the tension at the top must support the weight of the entire string. Substituting either condition, we obtain

$$T = \lambda gy.$$

Massless String wrapped around a Rough Pole

So far, we have been dealing with pulleys without friction between the axle and the rope. In the following problem, let us analyze a situation where a static massless rope is wrapped around a rough cylindrical pole.

Problem: A massless rope is wrapped around a rough pulley with a coefficient of static friction μ. The segment of the rope in contact with the pulley subtends an angle ϕ. If one end of the rope is vertically attached to a weight

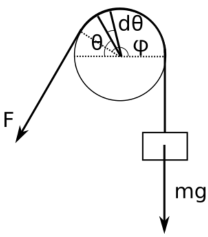


Figure 7.24: Holding a weight around a rough pole

of mass m, what is the minimum force F that has to be exerted at the other end for the rope and mass to remain static?

To solve this problem, we first note that the tension at the right end of the string segment in contact with the pulley must be mg, so that the weight remains stationary. Furthermore, the tension should vary along the string segment in contact with the pulley as friction can aid in preventing the weight from sliding downwards in the ideal scenario. The shape of the string segment is constrained to be an arc on a circle. Hence, infinitesimal segments of the string in contact with the pulley can be considered in polar coordinates.

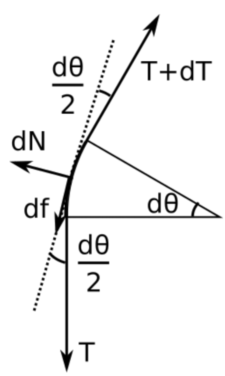


Figure 7.25: Infinitesimal sector of the pole and string

The forces that act on this infinitesimal segment of string that is between angles θ and $\theta + d\theta$, as measured clockwise from the horizontal, are depicted above. dN is the normal force, df is the friction force while T and $T + dT$ refer to the tensions at its ends at angular coordinates θ and $\theta + d\theta$. Balancing forces in the tangential direction,

$$(T + dT)\cos\frac{d\theta}{2} = T\cos\frac{d\theta}{2} + df.$$

Friction must be pointing in the anti-clockwise direction so that F (on the left) is a minimum. Furthermore, using the small angle approximation $\cos x \approx 1$ when x is small,

$$T + dT = T + df$$
$$dT = df.$$

Now, balancing forces in the radial direction,

$$dN = T\sin\frac{d\theta}{2} + (T + dT)\sin\frac{d\theta}{2}.$$

Using the small angle approximation $\sin x \approx x$ and discarding second-order infinitesimal terms,

$$dN = Td\theta.$$

Lastly, there is a constraint that

$$df \leq \mu dN$$
$$\implies dT = df \leq \mu T d\theta.$$

Shifting T to the left-hand side and integrating,

$$\int_F^{mg} \frac{1}{T} dT \leq \int_{\theta_0}^{\theta_0+\phi} \mu d\theta,$$

where θ_0 corresponds to the angle of the left end of the segment in contact with the pole. Then,

$$\ln\left|\frac{mg}{F}\right| \leq \mu\phi$$
$$F \geq mge^{-\mu\phi}. \tag{7.3}$$

It can be seen that the minimum amount of force required to maintain the system in static equilibrium decays exponentially with ϕ. Thus, friction can be leveraged to hold heavy objects (the above suggests that we should wrap more rounds of the rope around the pulley). Now, what is the minimum external force F' required for the weight to just begin to move upwards (impending motion)? In such a scenario, the impending motion is opposite to that in the previous situation as the mass now tends to move upwards. Hence, friction will act in the opposite direction and F' can be computed as

$$F' = mge^{\mu\phi}.$$

In the previous inequality, F' plays the role of mg while F is replaced by mg. Thus, friction is beneficial when we want to keep objects stationary, but detrimental when we need to move them. Actually, we can observe that the assumption of the rope taking the form of an arc was not essential in our treatment. Therefore, the above result can in fact be generalized to a rope clinging on a surface of an arbitrary shape (the angle ϕ just refers to the angle subtended by the normals from the two extreme points of the rope that are in contact with the surface).

7.5 Statically Indeterminate Situations

Hence far, the problems that we have encountered are all solvable based on the conditions for static equilibrium alone. In two-dimensional problems, applying the conditions for static equilibrium to a single system only results in three equations — namely, two for translational equilibrium and one for rotational equilibrium. Thus, static situations can be solved purely through these conditions if there are three or fewer independent force variables. However, in scenarios where the total number of force variables exceeds the number of degrees of freedom of a system — the number of translational and rotational motions that are independent — not all force variables can be completely solved for. This occurs when the system has too many supports and is known as a statically indeterminate system. The values of the forces in the final configuration depend on how the system was assembled in the first place.

Consider the following example: a block of mass m lies motionless on top of a table. The block is also vertically attached to the ceiling via a massless string. What is the normal force N on the block due to the table and the tension in the string, T?

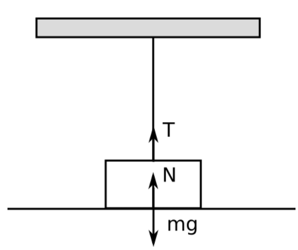


Figure 7.26: Block on table

Well, we obtain the lone equilibrium equation by balancing forces in the vertical direction:

$$T + N = mg.$$

Evidently, there are two unknowns but only one equation. The forces are indeterminate and their magnitudes depend on the way the system was constructed. In reality, bodies deform under stress and this affects the forces in a system, as we shall see. Let the final length of the string be l and consider the three different scenarios below.

In the first scenario, a rope of length $l - \epsilon_1$, where ϵ_1 is small, is tied to the ceiling. Then, the block is attached to the rope and allowed to come to rest, stretching the rope to length l in the process. At this point $T = mg$. Finally, the table is slid under the block. In this case, it is evident that

$$T = mg,$$
$$N = 0.$$

In a different situation, the block is first placed on the table. $N = mg$ at this juncture in order for the block to remain stationary. Then, a rope of length l is tied from the block to the ceiling. In the final configuration,

$$T = 0,$$
$$N = mg.$$

In the last situation, the block is first placed statically on the table. The magnitude of the normal force is still equal to the weight at this point. Now, a rope of length $l - \epsilon_2$ where $\epsilon_2 < \epsilon_1$ is attached to the block. The rope is then pulled upwards and tied to the ceiling, extending to a length l in the process. In this case, both the magnitudes of tension and the normal force acquire intermediate values between 0 and mg. Their exact values depend on the value of ϵ_2 and the elastic modulus of the rope.

Generally, deformations have to be considered and the rigid body assumption is discarded in order to generate more equations to solve a statically indeterminate problem. However, such situations are not particularly common — one should rather pay heed to identifying special conditions regarding the forces in a system that result from how the system was placed. Consider the following example.

Problem: Cylinder A, of radius R, is initially placed on the ground and gently rests against the wall. Now, cylinder B, of radius $r < R$, is placed on top of A, as shown in Fig. 7.27, while A is gently held by a person. Both cylinders remain static at this juncture. If the person now retracts his hand from cylinder A, what conditions must be fulfilled for the cylinders to remain in static equilibrium? The coefficients of static friction between cylinder A and the ground, the cylinders and the wall and between the cylinders are μ_1, μ_2 and μ_3 respectively. The cylinders both have mass m.

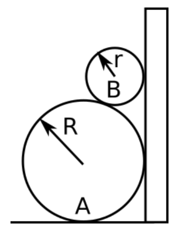


Figure 7.27: Two cylinders

As always, a free-body diagram for each of the two cylinders is drawn (Fig. 7.28).

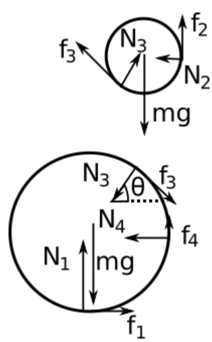


Figure 7.28: Free-body diagrams

There are a total of 8 variables to solve for. However, there are only 6 possible equilibrium equations relating them. Hence, there must be another constraint imposed on this system so that the problem is tractable. The critical observation is that when cylinder B is placed onto cylinder A, there is no tendency for cylinder A to move to the right. Hence, the normal force N_4 must be zero. Correspondingly, f_4 is also zero. Now, there are exactly 6 variables and 6 equations which enables the system to be solvable. Balancing torques about both cylinders about the axes through their centers,

$$f_1 = f_2 = f_3.$$

Hence, we shall just denote the friction forces as f henceforth. Balancing forces in the vertical and horizontal directions for both cylinders, we obtain the following set of simultaneous equations.

$$N_1 = mg + f\cos\theta + N_3\sin\theta, \tag{7.4}$$

$$f(1+\sin\theta) = N_3\cos\theta, \tag{7.5}$$

$$f(1+\cos\theta) + N_3\sin\theta = mg, \tag{7.6}$$

$$N_2 + f\sin\theta = N_3\cos\theta. \tag{7.7}$$

From Eqs. (7.5) and (7.7) (or from balancing forces on the combined system in the horizontal direction), it can immediately be seen that

$$N_2 = f.$$

Hence, for the system to remain static,

$$\left|\frac{f}{N_2}\right| \leq \mu_2$$
$$\implies \mu_2 \geq 1.$$

Next, f can be solved via Eqs. (7.5) and (7.6),

$$f = \frac{\cos\theta}{1+\sin\theta+\cos\theta}mg.$$

Substituting this into Eq. (7.5),

$$N_3 = \frac{1+\sin\theta}{1+\sin\theta+\cos\theta}mg.$$

Finally, substituting the relevant expressions for f and N_3 into Eq. (7.4),

$$N_1 = \frac{2+2\sin\theta+\cos\theta}{1+\sin\theta+\cos\theta}mg.$$

Hence, the conditions for the system to remain static are

$$\left|\frac{f}{N_1}\right| = \left|\frac{\cos\theta}{2+2\sin\theta+\cos\theta}\right| \leq \mu_1. \tag{7.8}$$

$$\left|\frac{f}{N_3}\right| = \left|\frac{\cos\theta}{1+\sin\theta}\right| \leq \mu_3. \tag{7.9}$$

Note that

$$\cos\theta = \frac{R-r}{R+r} > 0,$$
$$\sin\theta = \frac{2\sqrt{Rr}}{R+r} > 0.$$

Hence, the values in the absolute value brackets are both positive. Substituting these into Eq. (7.8),

$$(1-3\mu_1)R - 4\mu_1\sqrt{Rr} - (1+\mu_1)r \leq 0.$$

The left-hand side can be factorized into

$$(\sqrt{R}+\sqrt{r})[(1-3\mu_1)\sqrt{R} - (1+\mu_1)\sqrt{r}] \leq 0.$$

Since the term enclosed in brackets on the left is always positive,

$$\frac{\sqrt{r}}{\sqrt{R}} \geq \frac{1-3\mu_1}{1+\mu_1}.$$

Lastly, from Eq. (7.9),

$$(1-\mu_3)R - 2\mu_3\sqrt{Rr} - (1+\mu_3)r \leq 0$$

$$\left(\sqrt{R}+\sqrt{r}\right)\left[(1-\mu_3)\sqrt{R} - (1+\mu_3)\sqrt{r}\right] \leq 0$$

$$\frac{\sqrt{r}}{\sqrt{R}} \geq \frac{1-\mu_3}{1+\mu_3}.$$

Based on the relative magnitudes of μ_1 and μ_3, $\frac{\sqrt{r}}{\sqrt{R}}$ has to satisfy the stricter lower bound. Furthermore, recall that there is an upper bound that is given as $\frac{\sqrt{r}}{\sqrt{R}} < 1$.

7.6 Virtual Work

Recall that the infinitesimal work done by a force $\boldsymbol{F}$ on a point that undergoes a displacement $d\boldsymbol{r}$ is

$$dW = \boldsymbol{F} \cdot d\boldsymbol{r}.$$

In static systems, a concept of virtual work can be conjured by imagining a virtual displacement $\delta\boldsymbol{r}$ of a point at which a force $\boldsymbol{F}$ acts on. Recall that a general displacement $\delta\boldsymbol{r}$ of a point on a rigid body consists of both a translation and a rotation about a reference point.

$$\delta\boldsymbol{r} = \delta\boldsymbol{D}_{ref} + \delta\boldsymbol{\theta} \times \boldsymbol{r}_{ref}.$$

$\delta\boldsymbol{D}_{ref}$ is a small translation of a reference point O on the body in a certain direction (which drags the entire body with it while maintaining its

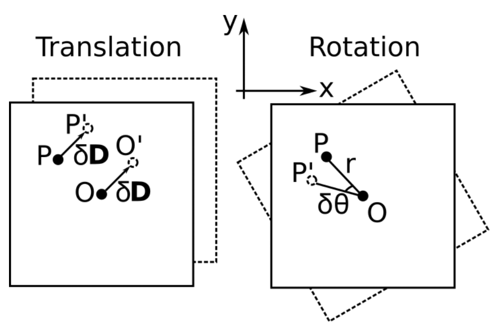


Figure 7.29: Translation and rotation about a point in two dimensions

orientation relative to O). $\delta\boldsymbol{\theta}$ denotes an infinitesimal rotation of the rigid body about point O of magnitude $\delta\theta$, in the plane perpendicular to $\delta\boldsymbol{\theta}$, and whose direction is given by the right-hand corkscrew rule. Note that even though general finite rotations cannot be represented by vectors, an infinitesimal rotation can be represented by a vector[3] (it can be seen as the angular velocity multiplied by a small time interval). $\boldsymbol{r}_{ref}$ is the position vector of the point of concern (whose displacement is $\delta\boldsymbol{r}$) on the body relative to point O. Correspondingly, the total virtual work done by a force on a point on a body that undergoes a virtual displacement $\delta\boldsymbol{r}$ is

$$\delta W = \boldsymbol{F} \cdot (\delta\boldsymbol{D}_{ref} + \delta\boldsymbol{\theta} \times \boldsymbol{r}_{ref}).$$

Using the triple-product rule

$$\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}).$$

The total virtual work can be expressed in a familiar form.

$$\delta W = \boldsymbol{F} \cdot \delta\boldsymbol{D}_{ref} + \delta\boldsymbol{\theta} \cdot (\boldsymbol{r}_{ref} \times \boldsymbol{F}).$$

The first term can be seen as the translational work done due to the force (associated with the virtual displacement of point O) while the second term can be understood as the rotational work done due to the torque produced by the force about the reference point O.

The total virtual work done by N forces at various points on a body due to a virtual displacement of the body is then given by

$$\delta W = \left(\sum_{i=1}^{N} \boldsymbol{F}_i\right) \cdot \delta\boldsymbol{D}_{ref} + \delta\boldsymbol{\theta} \cdot \sum_{i=1}^{N} \boldsymbol{\tau}_i.$$

Since internal forces result in no net force and torque, virtual work is only produced by external forces.

[3] The reason behind this is that three-dimensional rotations are not commutative (the order of rotation matters). Pick up your pen and hold it vertically between your thumb and index finger, with the tip pointing downwards. Now, rotate the pen by 90° such that the tip points towards you. Afterwards, rotate it by 90° again such that the tip points towards your palm. You will see that if you repeat these rotations in the reverse order, a different final configuration will be obtained (the pen tip points towards you). However, for infinitesimal rotations, this error is negligible and an infinitesimal rotation can be represented by a vector denoting the rotations about three different axes (without any mention of the sequence of rotations).

7.6.1 *The Principle of Virtual Work*

The principle of virtual work states that a system is in static equilibrium if and only if the total virtual work done by external forces acting on the system is zero for all virtual displacements of the system, consistent with its constraints.

Proof:

$\Longrightarrow$ If a system is in static equilibrium, the vector sum of forces is zero and the net torque about any point, including the reference point O, is zero — implying that $\sum \boldsymbol{F} = \mathbf{0}$ and $\sum \boldsymbol{\tau} = \mathbf{0}$. Hence,

$$\delta W = 0.$$

$\Longleftarrow$ Consider a pure virtual translation in the x-direction, $\delta \boldsymbol{D}_x$. Let F_{ix} denote the x-component of the ith force. Then, if the virtual work due to a virtual translation in the x-direction is zero,

$$\delta W = \left(\sum \boldsymbol{F}_i\right) \cdot \delta \boldsymbol{D}_x = \left(\sum F_{ix}\right) \delta D_x = 0$$

$$\Longrightarrow \sum F_{ix} = 0.$$

This implies that there must be no net force in the x-direction. A similar conclusion can be obtained for the other directions. Now, consider a pure virtual rotation of the body about the x-direction with respect to a certain origin. Then, if the total virtual work done due to this virtual rotation is zero,

$$\delta W = \hat{\boldsymbol{\imath}} \delta\theta \cdot \sum \boldsymbol{\tau} = 0$$

$$\Longrightarrow \sum \tau_x = 0.$$

Hence, the net torque about the origin in the direction of the infinitesimal rotation must be zero. Repeating this process for rotations in other directions, one can show that the system is in rotational equilibrium. Since the system must be in both translational and rotational equilibrium, it has attained a state of static equilibrium.

Applications

The principle of virtual work really shines in determining configurations of connected components that are possible for static equilibrium. Such utility stems from the fact that the virtual work of internal forces, which come in opposite and collinear pairs, can be neglected entirely. Hence, there is no need to "dismember" the connected components and to consider internal forces when applying the principle of virtual work.

Problem: Determine the angle θ for which the system below can remain in static equilibrium. The members are incompressible and have lengths l. The left member is pinned at one end while the right member is placed on a roller. An external force P is exerted on the frictionless pin that connects the two members. The relaxed length of the spring, with spring constant k, is l_0. Neglect the weight of the members.

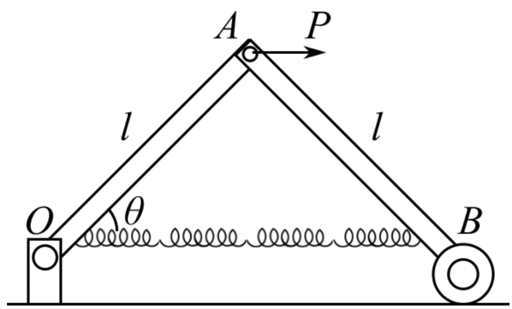


Figure 7.30: Connected members

Internal Forces: We shall first solve the problem by considering internal forces. We divide the entire system into three sub-systems — the left member, the right member and the pin connecting them. Drawing their free-body diagrams,

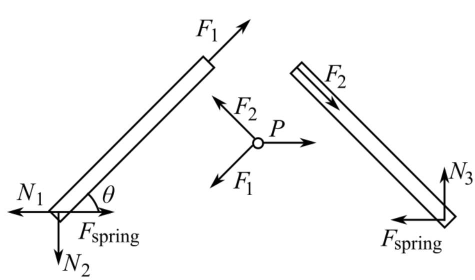


Figure 7.31: Free-body diagrams

N_1, N_2 and N_3 are forces on the members due to their supports at the bottom. As the left member is pinned, it can experience a force due to the pin in both the horizontal and vertical directions. Next, the forces F_1 and F_2 are exerted by the central pin on the respective members. They must act along the longitudinal directions of the members which are two-force members.

Balancing forces in the vertical and horizontal directions of the pin,

$$F_1 \sin\theta = F_2 \sin\theta$$
$$\implies F_1 = F_2 = F,$$

where we introduce a new unified variable F.

$$P = 2F\cos\theta$$
$$F = \frac{P}{2\cos\theta}.$$

Balancing forces on the right member in the horizontal direction,

$$F\cos\theta = F_{spring}.$$

The force due to the spring can be computed as

$$F_{spring} = k(2l\cos\theta - l_0).$$

Hence,

$$\frac{P}{2} = k(2l\cos\theta - l_0)$$
$$\cos\theta = \frac{P}{4kl} + \frac{l_0}{2l}.$$

Principle of Virtual Work: Let us now solve the problem using virtual work. Define the x-axis to be along the horizontal direction with the origin at O. We shall consider the virtual work done by the external forces when the angle θ increases by a small angle $\delta\theta$. Let the x-coordinates of point A and B be x_A and x_B respectively.

$$x_A = l\cos\theta,$$
$$x_B = 2l\cos\theta.$$

The only external forces that act on points which undergo virtual displacements when θ is increased by $\delta\theta$ are the force P at point A and the spring force on point B. Since the external forces only act in the horizontal direction, we simply need to consider the virtual displacements of A and B in the horizontal direction in computing virtual work. These virtual displacements, δx_A and δx_B, can be computed via the derivative of their respective x-coordinates with respect to θ.

$$\delta x_A = -l \sin\theta \delta\theta,$$

$$\delta x_B = -2l \sin\theta \delta\theta.$$

The total virtual work done by the forces on the system is then

$$\delta W = \delta W_P + \delta W_{spring} = -l \sin\theta \delta\theta P + 2l \sin\theta \delta\theta F_{spring} = 0.$$

Substituting the expression for F_{spring}, we obtain the same result:

$$\cos\theta = \frac{P}{4kl} + \frac{l_0}{2l}.$$

It can be seen that there was completely no need to consider internal forces and external forces that produce zero work when the system undergoes a certain virtual displacement. Hence, the principle of virtual work can greatly simplify one's analysis in appropriate situations.

7.6.2 *Potential Energy*

If a virtual displacement leads to work done by only conservative forces, the principle of virtual work can be expressed in an equivalent form. Let q_i represent a generalized coordinate of the system — it could be an angular or translational coordinate. In static systems, the number of independent coordinates required to define the state of a system is equal to the number of degrees of freedom, f. Following from this, the potential energy of a system can be defined by f coordinates. Then, the virtual work done by conservative forces when the system undergoes a virtual displacement from a particular coordinate q_1 to $q_1 + dq_1$ is

$$\delta W_{cons} = -(U(q_1 + dq_1, q_2, \ldots) - U(q_1, q_2, \ldots)),$$

where U refers to the total potential energy of the system. By the principle of virtual work, if the work done is only due to conservative forces, then $\delta W_{cons} = 0$,

$$\implies U(q_1 + dq_1, q_2, \ldots) - U(q_1, q_2, \ldots) = 0.$$

If we divide the above expression by dq_1 and take the limit as $dq_1 \to 0$,

$$\frac{\partial U}{\partial q_1} = 0.$$

This means that at positions of static equilibrium, the total potential energy of a system should undertake a stationary value with respect to a coordinate q if virtual displacements in coordinate q result in virtual work done by only conservative forces. This provides another handy tool in enforcing static equilibrium for a system that is under the influence of gravitational forces, spring forces and other conservative forces. It is still possible to apply this technique to systems with non-conservative forces if the virtual displacements are set to be perpendicular to these forces. Systems with normal forces immediately come to mind.

Problem: A uniform ladder of mass m and length l is placed on a frictionless ground and leans on a frictionless wall. There are two springs, of spring constants k_1 and k_2, attached to the left end and the center of the ladder respectively. The other ends of both springs are both attached to the left end of the ceiling. If the height between the ceiling and the ground is h, determine the angles θ at which the ladder is in static equilibrium. Assume that the relaxed lengths of the springs are zero.

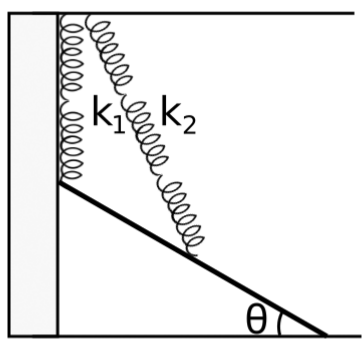


Figure 7.32: Leaning ladder

The total potential energy of the system is

$$U = mg\frac{l}{2}\sin\theta + \frac{1}{2}k_1(h - l\sin\theta)^2 + \frac{1}{2}k_2\left[\left(h - \frac{l}{2}\sin\theta\right)^2 + \frac{l^2\cos^2\theta}{4}\right],$$

$$\frac{dU}{d\theta} = mg\frac{l}{2}\cos\theta - k_1(h - l\sin\theta)l\cos\theta - \frac{k_2hl}{2}\cos\theta$$

$$= \frac{\cos\theta l}{2}\left[mg - 2k_1(h - l\sin\theta) - k_2h\right].$$

When the system is at static equilibrium, $\frac{dU}{d\theta} = 0$

$$\implies \cos\theta(2k_1l\sin\theta + mg - 2k_1h - k_2h) = 0.$$

The two solutions to this equation are

$$\cos\theta = 0$$
$$\implies \theta = \frac{\pi}{2}$$

and

$$\sin\theta = \frac{2k_1h + k_2h - mg}{2k_1l}$$
$$\theta = \sin^{-1}\left(\frac{2k_1h + k_2h - mg}{2k_1l}\right)$$

if $0 \leq \frac{2k_1h+k_2h-mg}{2k_1l} \leq 1$. Note that θ is constrained to be between 0 and $\frac{\pi}{2}$. $\theta > \frac{\pi}{2}$ would imply that the ladder is on the left of the wall which is physically impossible. The equations above could have also been obtained by considering torques about the point of intersection of the normal forces due to the wall and the ground.

Problem: Two identical members, of length l and mass m, are fixed perpendicularly with respect to each other. The two free ends of the members are attached to springs of spring constant $k = \frac{mg}{l}$ that are vertically connected to the ceiling. If the relaxed length of the springs is l_0, what are the possible configurations of the system such that it is in static equilibrium?

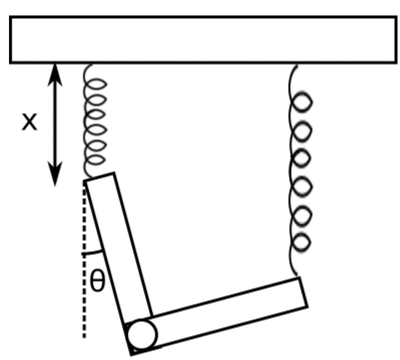


Figure 7.33: Hanging members

The system has two degrees of freedom. Hence, we need two coordinates to uniquely define its state. We will adopt the coordinates x and θ as labeled in Fig. 7.33. The total potential energy of the system is

$$U = -mg\left(x + \frac{l}{2}\cos\theta\right) - mg\left(x + l\cos\theta - \frac{l}{2}\sin\theta\right)$$
$$+\frac{1}{2}k(x - l_0)^2 + \frac{1}{2}k(x + l\cos\theta - l\sin\theta - l_0)^2.$$

The partial derivatives of the total potential energy with respect to both x and θ must be zero for the system to remain in static equilibrium.

$$\frac{\partial U}{\partial x} = -2mg + k(x - l_0) + k(x + l\cos\theta - l\sin\theta - l_0) = 0, \quad (7.10)$$
$$\frac{\partial U}{\partial \theta} = mgl\left(\frac{3}{2}\sin\theta + \frac{1}{2}\cos\theta\right)$$
$$-kl(x + l\cos\theta - l\sin\theta - l_0)(\sin\theta + \cos\theta) = 0. \quad (7.11)$$

Multiplying Eq. (7.10) by $\sin\theta + \cos\theta$ and adding it to Eq. (7.11), divided by l, yields the following after some rearrangement.

$$k(x - l_0) = \frac{mg(\sin\theta + 3\cos\theta)}{2(\sin\theta + \cos\theta)} \implies x = \frac{\sin\theta + 3\cos\theta}{2(\sin\theta + \cos\theta)}l + l_0,$$

since $k = \frac{mg}{l}$. Substituting this expression into Eq. (7.10),

$$(\cos\theta - \sin\theta)(\cos\theta + \sin\theta + 1) = 0.$$

One solution is

$$\tan\theta = 1$$
$$\implies \theta = \frac{\pi}{4},$$
$$x = \frac{mg}{k} + l_0 = l + l_0.$$

Another solution is

$$\sin\theta + \cos\theta = -1$$
$$\sqrt{2}\sin\left(\theta + \frac{\pi}{4}\right) = -1$$
$$\theta = -\frac{\pi}{2} \quad \text{or } \pi,$$
$$x = \frac{l}{2} + l_0 \quad \text{or} \quad \frac{3l}{2} + l_0.$$

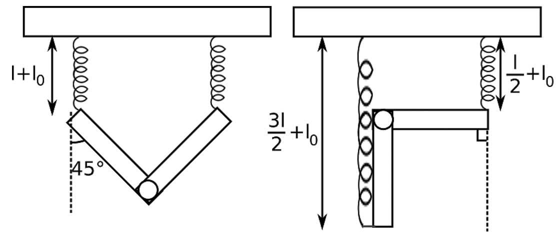


Figure 7.34: Static equilibrium positions

The two possible configurations are depicted in the figure above. The two cases $\theta = -\frac{\pi}{2}$ and $\theta = \pi$ are combined into the diagram on the right, as the two situations only differ by a flip in the horizontal direction.

7.6.3 *Stability of Equilibrium*

The potential energy function of a conservative system in terms of a certain coordinate can be used to investigate the stability of an equilibrium with respect to that coordinate. We will only analyze systems with a single degree of freedom. Recall that at points of equilibrium, the first derivative of the potential energy with respect to a coordinate is zero, that is, $\frac{dU}{dq} = 0$.

- In a stable equilibrium with respect to a certain coordinate, any deviation in that particular coordinate tends to produce a force that minimizes the deviation and returns the system back to the equilibrium position. As a conservative force is oriented towards points of lower potential energy, this requires the potential energy function to be a local minimum, $\frac{d^2U}{dq^2} > 0$. This is commensurate with our intuition that a lower potential energy implies a more stable state. A possible configuration is shown in the left-most diagram in Fig. 7.35. A circle is attached to a ceiling like a pendulum. Any deviation in its angle with respect to the vertical will be reduced by the torque arising from its weight.
- In a neutral equilibrium with respect to a coordinate, any deviation in that particular coordinate does not lead to a response by the conservative system to amplify or reduce the deviation. A truly neutral equilibrium requires all (higher-order) derivatives of the potential energy function with respect to that coordinate to be zero. In a neutral equilibrium, the system possesses a constant potential energy with respect to a particular coordinate, though the neutral state may have a finite width. An example of a neutral equilibrium would be a cylinder lying on a tabletop, as shown in the middle diagram of Fig. 7.35. If the cylinder translates forwards slightly, there is no tendency for further translation backwards or forwards.

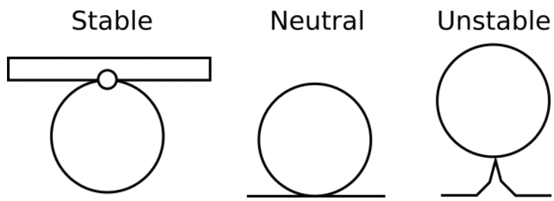


Figure 7.35: Illustrations of different stabilities

- In an unstable equilibrium with respect to a certain coordinate, any deviation in that particular coordinate tends to produce a force that further amplifies the deviation. This requires the potential energy to be a local maximum relative to that coordinate, $\frac{d^2U}{dq^2} < 0$. A example of an unstable equilibrium would be the cylinder resting on top of a small hill, as shown in the rightmost diagram in Fig. 7.35. A slight rotation of the cylinder with respect to its point of contact with the hill will produce a gravitational torque that causes the cylinder to rotate further away from its equilibrium position.

In the rare case where the second derivative is zero, higher order derivatives of U need to be examined. Performing a Taylor expansion of $U(q)$ about the equilibrium position q_0,

$$U(q) = U(q_0) + \frac{U''(q_0)}{2!}(q-q_0)^2 + \frac{U'''(q_0)}{3!}(q-q_0)^3 + \cdots + \frac{U^n(q_0)}{n!}(q-q_0)^n + \cdots .$$

Note that we have omitted the term of degree one, as $U'(q_0) = 0$ by definition of an equilibrium point. Now if the first non-zero derivative is of odd order, the equilibrium point is neither a maximum nor minimum as increasing and decreasing $q - q_0$ changes the value of $U(q)$ in opposite directions. However, if the first non-zero derivative is of even order, the equilibrium point is a turning point. If the first non-zero derivative is of even order and has a positive value, the equilibrium is stable (as adjacent values of U are larger). Otherwise, if it is of even order and negative, it is unstable.

Problem: An equilateral triangle, of side length l and uniform mass density, is at static equilibrium when it is placed in a gap of width d. One side of the triangle is parallel to the horizontal such that the axis of the triangle is placed vertically in the gap. Determine the minimum value of d for which the triangle remains at stable equilibrium. (Estonian Finnish Olympiad)

The only degree of freedom of the triangle is a rotation about the vertical axis. Our strategy would be to determine the coordinates of the vertex of the triangle in the gap and use the fact that the CG is a distance $\frac{l}{\sqrt{3}}$ away

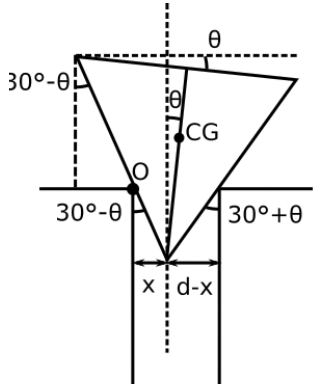


Figure 7.36: A slight displacement

from any vertex of an equilateral triangle.[4] Let the origin be defined at the left end of gap, denoted as O in the figure. Let the coordinates of the bottom tip be (x, y). Then,

$$\frac{x}{-y} = \tan(30^\circ - \theta),$$
$$\frac{d-x}{-y} = \tan(30^\circ + \theta)$$
$$y = -\frac{d}{\tan(30^\circ + \theta) + \tan(30^\circ - \theta)}.$$

The y-coordinate of the center of mass y_{CG} is the addition of the vertical component of the length between the vertex and the centroid, to y.

$$y_{CG} = -\frac{d}{\tan(30^\circ + \theta) + \tan(30^\circ - \theta)} + \frac{l}{\sqrt{3}} \cos\theta.$$

Since the gravitational force is the only conservative force in this conservative system, the second derivative of y_{CG} at $\theta = 0$ must be greater than zero for the system to be in stable equilibrium at $\theta = 0$.

$$\frac{dy_{CG}}{d\theta} = \frac{d}{[\tan(30^\circ + \theta) + \tan(30^\circ - \theta)]^2} \left[\sec^2(30^\circ + \theta) - \sec^2(30^\circ - \theta)\right] - \frac{l}{\sqrt{3}} \sin\theta.$$

[4]Note that the center of mass of a triangle is the point of concurrency of the medians of the triangle. Furthermore, the point of concurrency divides each median into a 2 : 1 ratio, with the segment closer to the base having a shorter length.

As expected, substituting $\theta = 0$ into the first derivative gives zero. The second derivative is

$$\begin{aligned}\frac{d^2 y_{CG}}{d\theta^2} = &-\frac{2d}{[\tan(30^\circ+\theta)+\tan(30^\circ-\theta)]^3}[\sec^2(30^\circ+\theta)\\ &-\sec^2(30^\circ-\theta)]^2 - \frac{l}{\sqrt{3}}\cos\theta\\ &+\frac{2d}{[\tan(30^\circ+\theta)+\tan(30^\circ-\theta)]^2}[\sec^2(30^\circ+\theta)\tan(30^\circ+\theta)\\ &+\sec^2(30^\circ-\theta)\tan(30^\circ-\theta)].\end{aligned}$$

Substituting $\theta = 0$ and concluding that the second derivative at this point must be greater than zero,

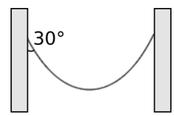
$$\begin{aligned}\frac{4}{\sqrt{3}}d - \frac{l}{\sqrt{3}} &> 0\\ d &> \frac{l}{4}.\end{aligned}$$

Problems

Balancing Forces and Torques

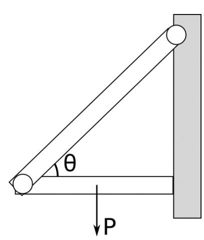
1. *Hanging Chain**

A uniform chain of mass m is hung from two walls with ends at equal heights. If the slope of the chain makes an angle of $30°$ with the vertical at its ends, determine the tension T at the center of the chain.



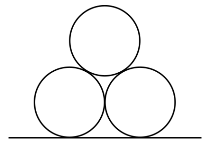
2. *Maintaining a Member**

Two massless members are connected by frictionless pins to each other as shown in the figure below. One member is attached to the wall using a pinned support while the other member is horizontal and touches the wall. A force P is exerted at the center of the bottom member. If the angle between the two members is $\theta > 0$ and if the coefficient of static friction between the bottom member and the wall is μ, what condition must θ satisfy for the system to attain static equilibrium?



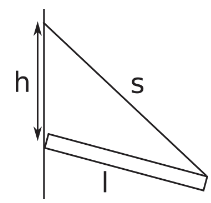
3. *Three Circles**

Three identical circles, of uniform mass m, are arranged as shown in the figure below. If the coefficients of static friction between a circle and the ground and between the circles are μ_1 and μ_2 respectively, determine the conditions for the system to remain in static equilibrium.



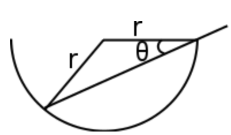
4. *Hanging Rod**

A rod, of uniform mass density and length l, is connected to a frictionless wall via a massless string of length s. Determine height h, which is defined in the figure below, for which the rod can remain in static equilibrium.



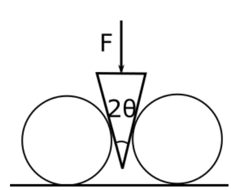
5. *Resting Rod***

A rod of length l rests in a massive semi-circular bowl of radius r with its two ends at different halves of the bowl. It makes an angle θ with the horizontal. Assuming that all surfaces are frictionless and that the rod protrudes out of the bowl, determine θ (if any) required for the rod to remain in static equilibrium and the conditions for such θ to exist.



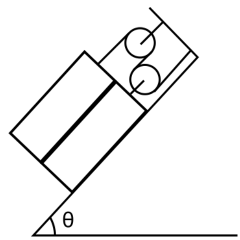
6. *Driving a Wedge***

An isosceles wedge of angle 2θ and mass m is sandwiched between two circles of mass M. A slowly increasing force F is exerted on the wedge. If the coefficients of static friction between the circles and the wedge and between the ground and the circles are μ_0 and μ respectively, what is the condition on μ_0 for the system to remain at static equilibrium before a critical value of F is reached? What is the critical value of F for which one can drive the wedge deeper between the circles?



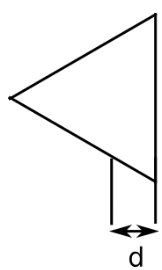
7. *Two Bricks**

Two identical bricks of mass m are attached to two pulleys and placed on an inclined plane, as shown in the figure below. The pulleys are connected to a massive wall. If the coefficients of static friction between the brick and the ramp and between the bricks are μ_1 and μ_2 respectively, determine the angle θ for which the bricks begin to slip if θ is gradually increased from 0.



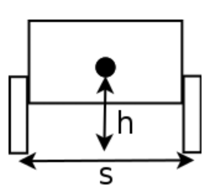
8. *Supporting a Wedge**

An equilateral triangle, of length l and uniform mass density, is supported by a gap as shown below. Determine the minimum value of d for the wedge to remain in static equilibrium. Assume all surfaces to be frictionless. (Estonian-Finnish Olympiad)



9. *Turning Car**

A car can be modeled as two wheels separated by a distance s, connected by a body in the middle. If the center of mass of the car is a height h above the ground and undergoes circular motion of radius l, determine the maximum angular velocity of the car's motion ω such that it does not topple. Assume that the coefficient of static friction between the wheels and the ground is large.

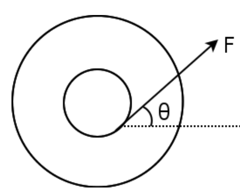


10. *Crawling on Ladder***

An ant of mass m is initially stationary at the top of a ladder of mass M and length l that leans against a smooth wall and rests on smooth ground. The ladder is initially held by you such that it is stationary. Now, you release your grip and the ant begins to travel down the ladder, towards the end at the ground. If the ladder remains stationary throughout the ant's motion, determine the time required by the ant to reach the bottom end.

11. *A Spool***

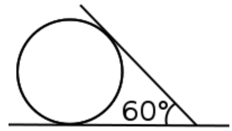
A spool, of total mass m, consists of an axle of radius r and a larger circle of radius R. A string is attached to the axle and pulled with a force F at an angle θ with respect to the horizontal. The spool lies on rough ground. ("Introduction to Classical Mechanics")



- In terms of r and R, what should angle θ be such that the spool remains in static equilibrium? Assume that static friction is large enough for the spool to remain static.
- Given r, R and the coefficient of static friction between the ground and the spool μ, what is the maximum value of F for which the spool does not move?
- Given R and μ, what should r be so that one can make the spool slip with the smallest F possible?

12. *Circle and Stick***

A stick of length l and mass m rests on a circle of mass $3m$, with a quarter of its length protruding above the circle as shown in the figure on the next page. The stick and the circle lie on top of a rough, horizontal table. If the stick subtends a $60°$ angle relative to the horizontal and friction is present between all surfaces, determine the conditions required for this system to remain in static equilibrium.



Strings

13. ***Tension in a String*****

A string with mass density λ and one end at x-coordinate x_0 is placed on an arbitrary frictionless surface described by $y(x)$, where the x-axis has been defined as the horizontal axis. If the tension at $x = x_0$ is T_0, determine $T(x)$. Do this in two ways: (1) by considering infinitesimal segments of string and (2) by considering virtual displacements. How would $T(x)$ vary if you fix its two ends and remove the surface such that the string now hangs under its own weight to form a new shape $y(x)$? The tension in the end at $x = x_0$ is still T_0.

14. ***Shape of String*****

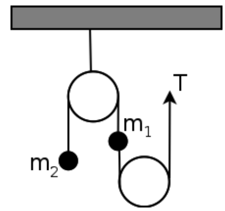
A massless string hangs between two walls at horizontal coordinates $x = 0$ and $x = l$. If an infinitesimal segment of string between horizontal coordinates x and $x + dx$ experiences a force $\lambda x dx$ vertically downwards, show that the shape of the string at equilibrium obeys

$$y = \frac{\lambda x^3}{6F} + y'(0)x + y(0)$$

where y is the vertical coordinate (positive upwards) and F is a constant. $y'(0)$ and $y(0)$ are the gradient and vertical coordinate of the end at $x = 0$ respectively. In your derivation, interpret the meaning of F.

15. ***Static Atwood*****

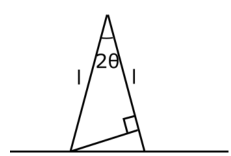
Consider the Atwood's machine on the next page; the strings subtend an angle θ with respect to the pulleys, with a coefficient of static friction μ. What are the minimum and maximum tensions that you can exert on the right end of the string such that the system remains static? Assume that $\frac{mg}{e^{\mu\theta}} > m_1 g$.



Virtual Work and Potential Energy

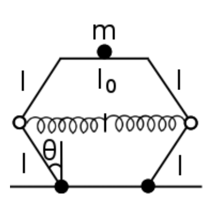
16. *String and Sticks***

Two identical uniform sticks of mass m and length l are pinned together. They both make an angle θ with respect to the vertical. A massless string is connected to the bottom of one rod and perpendicularly to the other rod as shown in the figure below. Determine the tension in the string if the table, on which the sticks are located, is frictionless. Solve this problem by balancing forces and torques, and applying the principle of virtual work.



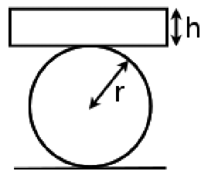
17. *Table***

The legs of a table are each formed by two massless rods of length l, connected by a pinned connection. The pinned connections are connected via a spring of rest length l_0 and spring constant k. If the base of the table is a massless rod of length l_0 and if the legs of the table are fixed to the ground, determine the possible angles of θ such that the system can remain in static equilibrium when a point mass m is placed at the center of the table. Examine the stability of the equilibrium(s).



18. *Block on Cylinder*****

A block of height h rests directly on top of a cylinder of radius r, whose cylindrical axis is in the plane of the table. If the block does not slip with respect to the cylinder, determine the condition for the block to be in a stable equilibrium.



Solutions

1. Hanging Chain*

Divide the chain into half about the midpoint and consider the free-body diagram of one of the halves. The tension at the bottom end of half the chain, which is at the center of the original chain, T, only points in the horizontal direction. Let the tension at the top of half the chain be T_0. It makes an angle of 30° with the vertical. The weight of this half section is $\frac{mg}{2}$. Hence, for static equilibrium to be attained,

$$\frac{mg}{2} = T_0 \cos 30^\circ$$
$$T_0 = \frac{mg}{\sqrt{3}},$$
$$T = T_0 \sin 30^\circ = \frac{mg}{2\sqrt{3}}.$$

2. Maintaining a Member*

Observe that the diagonal member is a two-force member. Hence, the force on that member can only be in the longitudinal direction. Correspondingly, the force due to that member on the horizontal member also makes an angle θ with respect to the horizontal. The free-body diagram is shown below.

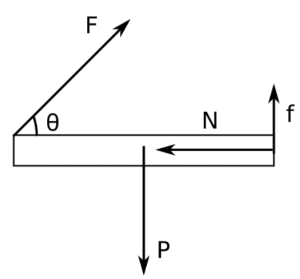


Figure 7.37: Free-body diagram

Taking torques about the right end, it can be deduced that

$$F \sin\theta = \frac{P}{2}$$
$$F = \frac{P}{2\sin\theta}.$$

Balancing forces,

$$f = P - F\sin\theta = \frac{P}{2},$$
$$N = F\cos\theta = \frac{P\cos\theta}{2\sin\theta},$$
$$\left|\frac{f}{N}\right| = |\tan\theta| \leq \mu.$$

It is obvious that the physical situation corresponds to $0 < \theta < \frac{\pi}{2}$. Since $\tan\theta$ is positive in this regime, we can simply remove the absolute value brackets to obtain

$$\tan\theta \leq \mu.$$

3. Three Circles*

Let the normal forces between a circle and the ground and between the bottom circles and the top circle be N_1 and N_2 respectively. Then, let the friction between the ground and the bottom left circle be f rightwards. The friction between the bottom left circle and the top circle is also f, in order for torques to be balanced about the bottom left circle. A similar situation occurs for the bottom right circle. Thus, all friction forces are of magnitude f. Balancing horizontal forces on the left circle,

$$N_2\cos 60^\circ = f(1 + \cos 30^\circ)$$
$$\frac{f}{N_2} = \frac{1}{2+\sqrt{3}} = 2 - \sqrt{3} \leq \mu_2.$$

By balancing forces on the top circle in the vertical direction,

$$2(N_2\sin 60^\circ + f\sin 30^\circ) = mg$$
$$f = \frac{mg}{4+2\sqrt{3}} = \frac{(2-\sqrt{3})mg}{2}.$$

Furthermore, by considering the three circles as an entire system,

$$N_1 = \frac{3mg}{2}$$

by symmetry. Hence,

$$\frac{f}{N_1} = \frac{2-\sqrt{3}}{3} \leq \mu_1.$$

4. Hanging Rod*

Observe that there are three forces on the rod — the tension force, normal force and its weight. Hence, in order for it to attain a state of static equilibrium, the lines of action of these forces must be concurrent.

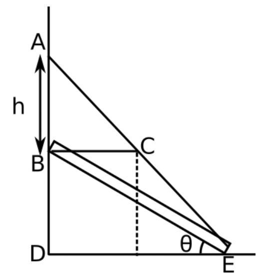


Figure 7.38: Lines of action

The lines of action of the normal force and the tension force intersect at point C. In order for the line of action of the weight to intersect at this point as well, segment BC must have length $\frac{l\cos\theta}{2}$. Observe that triangles ABC and ADE are similar, so that

$$\triangle ABC \sim \triangle ADE,$$

$$\frac{\overline{AB}}{\overline{AD}} = \frac{\overline{BC}}{\overline{DE}} \implies \frac{h}{\sqrt{s^2 - l^2\cos^2\theta}} = \frac{1}{2}.$$

Since, $h = \sqrt{s^2 - l^2\cos^2\theta} - l\sin\theta$,

$$2l\sin\theta = \sqrt{s^2 - l^2\cos^2\theta}$$

$$3l^2\sin^2\theta = s^2 - l^2$$

$$l^2\cos^2\theta = \frac{4l^2 - s^2}{3}$$

$$h = \frac{\sqrt{s^2 - l^2\cos^2\theta}}{2} = \frac{\sqrt{\frac{4s^2-4l^2}{3}}}{2} = \sqrt{\frac{s^2 - l^2}{3}}.$$

5. Resting Rod**

There are forces at three points of the rod — two normal forces and its weight.

Refering to the figure on the next page if we define the origin O to be at the left end of the rod, the x-coordinate of the point of intersection of lines

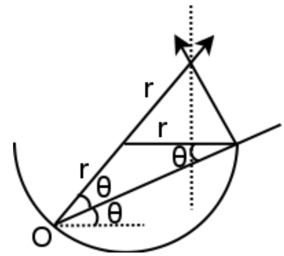


Figure 7.39: Resting rod

of action of the two normal forces is $2r\cos 2\theta$. Hence, for the weight and the two normal forces to be concurrent,

$$2r\cos 2\theta = \frac{l}{2}\cos\theta$$

$$\frac{8r}{l}\cos^2\theta - \cos\theta - \frac{4r}{l} = 0$$

$$\cos\theta = \frac{l + \sqrt{l^2 + 128r^2}}{16r},$$

where we have rejected the negative solution which implies that the two ends of the rod lie in the same half of the bowl. For θ to exist, the right-hand side must be smaller than or equal to one.

$$\frac{l + \sqrt{l^2 + 128r^2}}{16r} \leq 1$$

$$\sqrt{l^2 + 128r^2} \leq 16r - l$$

$$l^2 + 128r^2 \leq 256r^2 - 32rl + l^2$$

$$\implies r \geq \frac{l}{4}.$$

6. Driving a Wedge**

We first draw free-body diagrams of the wedge and circles in Fig. 7.40.

The friction forces on each circle must be the same to satisfy the equality of torques about the center. Furthermore, the friction forces on the bottoms of the circles must be equal for the horizontal forces on the entire system to be balanced. We then denote all friction forces as f. Balancing forces in the horizontal direction of a circle,

$$N_2\cos\theta = f(1 + \sin\theta).$$

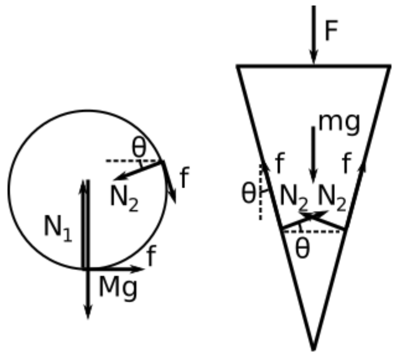


Figure 7.40: Free-body diagrams

In order for the system to remain in static equilibrium,

$$\frac{f}{N_2} = \frac{\cos\theta}{1+\sin\theta} \le \mu_0.$$

Note that f is always positive in this case, as N_2, $\cos\theta$ and $(1+\sin\theta)$ are non-negative for $0 \le \theta < \frac{\pi}{2}$. Thus, we did not need to consider the absolute value in $|f| \le \mu_0 N_2$. Balancing forces on the wedge in the vertical direction,

$$2(f\cos\theta + N_2\sin\theta) = F + mg.$$

Substituting the expression of N_2 in terms of f, we solve the equations to obtain

$$f = \frac{(F+mg)\cos\theta}{2(\sin\theta+1)},$$
$$N_2 = \frac{F+mg}{2}.$$

In fact, this is a general trend in problems in which an object is balanced between two identical circles, with friction between the object and the circles, and between the circles and the ground. The normal force that is exerted on a circle due to the object is half the total vertical force experienced by the object (including its weight), excluding the forces due to the circle. This can be seen from the fact that the zero net torque and force equations could have been rewritten for any general set-up. Next, if we consider the entire set-up as a whole, $2N_1$ must balance the entire weight of the system plus F.

Hence

$$N_1 = \frac{F + mg}{2} + Mg.$$

When the system is just about to slip, $f = \mu N_1$. Equating these two expressions and solving for F, impending motion occurs when F attains the value

$$F = \frac{mg(\mu \sin\theta + \mu - \cos\theta) + 2\mu Mg(\sin\theta + 1)}{\cos\theta - \mu\sin\theta - \mu}.$$

7. Two Bricks**

Let the friction forces between the block and the ramp and between the two blocks be f_1 and f_2 respectively. Similarly, define N_1 and N_2 as the normal force on the bottom block due to the ramp and that on the top block due to the bottom block. The key observation is that these bricks tend to slip simultaneously in opposite directions as they are connected by a fixed pulley. Assume that f_2 and f_1 are directed upwards and downwards along the ramp, respectively. In other words, the top block tends to slip downwards while the bottom block tends to slip upwards. Let the tension in the string, connected to the top block, be T and consider the forces on the block on top. At equilibrium,

$$T + f_2 = mg\sin\theta,$$
$$N_2 = mg\cos\theta.$$

When the top brick is about to slip, $f_2 = \mu_2 N_2$.

$$T + \mu_2 mg\cos\theta = mg\sin\theta.$$

Considering the forces on the two blocks as a whole,

$$3T = f_1 + 2mg\sin\theta,$$
$$N_1 = 2mg\cos\theta.$$

Similarly, at this juncture, $f_1 = \mu_1 N_1$.

$$3T = 2\mu_1 mg\cos\theta + 2mg\sin\theta.$$

Eliminating the tensions,

$$2\mu_1 mg\cos\theta + 2mg\sin\theta = 3mg\sin\theta - 3\mu_2 mg\cos\theta$$
$$\tan\theta = 2\mu_1 + 3\mu_2.$$

If the friction forces were assumed to point in the opposite directions, a negative value of θ will be obtained and hence, this case is rejected.

8. Supporting a Wedge**

Firstly, observe that the distribution of the normal force on the right side of the wedge generally depends on the width of the gap. When the gap attains its minimum width, we argue that the entire normal force on the right side is at the bottom tip of the wedge.

To understand this, consider torques about the bottom end of the right edge, B. The clockwise torque due to the normal force N_2 on the slanted side of the triangle balances the anti-clockwise torques due to the wedge's weight and possibly due to the normal force on the right edge. Observe that N_2 has a fixed magnitude for different d's as its vertical component must balance the wedge's weight. Therefore, as d decreases, the clockwise torque due to N_2 decreases as the length of the moment's arm is reduced — implying that the anti-clockwise torque due to the normal force on the right edge follows suit. In the boundary case where d reaches its minimum value, the normal force on the right edge must contribute zero anti-clockwise torque and act entirely at point B. Any gap of a smaller width would require a normal force outside the wedge, which is impossible.

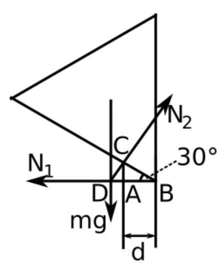


Figure 7.41: Three forces on the wedge

At the minimum gap width, the forces act on the wedge at exactly three points. Hence, the concurrency criterion can be used to determine the gap width. The lines of action of the forces are depicted above. The length of line segment AB is

$$l_{AB} = d.$$

Hence, the length of segment BC is

$$l_{BC} = \frac{l_{AB}}{\cos 30^\circ} = \frac{2d}{\sqrt{3}}.$$

The length of segment BD is

$$l_{BD} = \frac{l_{BD}}{\cos 30^\circ} = \frac{4d}{3}.$$

In order for the lines of action to coincide, l_{BD} must be equal to the horizontal distance between the center of mass of the wedge and the right side of the wedge, which is $\frac{l}{2\sqrt{3}}$. Thus,

$$\frac{4d}{3} = \frac{l}{2\sqrt{3}}$$
$$\implies d = \frac{\sqrt{3}l}{8}.$$

9. Turning Car**

Friction provides the required centripetal force for the car to undergo circular motion. Let the friction forces on the left and right wheels be f_1 and f_2 respectively, pointing towards the left. Let the corresponding normal forces be N_1 and N_2. For the car to remain in circular motion,

$$f_1 + f_2 = ml\omega^2,$$

where m is the mass of the car. Balancing forces in the vertical directions,

$$N_1 + N_2 = mg.$$

Now, balancing torques about the center of mass of the car,

$$(N_2 - N_1)\frac{s}{2} = (f_1 + f_2)h.$$

Note that it is essential to define our pivot at the center of mass as the body is accelerating. The net torque is not necessarily zero with respect to other origins. Solving this system of equations for N_1 and N_2,

$$N_2 = \frac{mg}{2} + \frac{ml\omega^2 h}{s},$$
$$N_1 = \frac{mg}{2} - \frac{ml\omega^2 h}{s}.$$

Since N_1 must be non-negative,

$$\omega \leq \sqrt{\frac{gs}{2lh}}.$$

10. Crawling on Ladder**

Consider the situation when the ant is at a distance x from the top of a ladder. In order for the ant to remain on the ladder, there must be a normal force on it due to the ladder that balances the $mg\cos\theta$ component of its weight perpendicular to the ladder. Generally, there is also a force f between the ladder and the ant in the direction parallel to the ladder, as depicted in the free-body diagrams below.

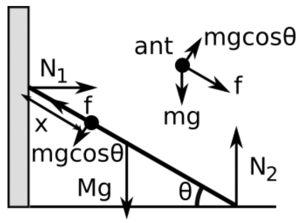


Figure 7.42: Free-body diagrams of ladder and ant

Balancing torques on the ladder about the point of intersection of the normal forces on the ladder due to the wall and the ground,

$$mg\cos\theta(l\cos^2\theta - x) = fl\sin\theta\cos\theta$$
$$f = \frac{mg(l\cos^2\theta - x)}{l\sin\theta}.$$

The acceleration of the particle along the ladder is due to f and the tangential component of its weight, $mg\sin\theta$.

$$f + mg\sin\theta = m\ddot{x}$$
$$\ddot{x} = \frac{g}{l\sin\theta}(l - x).$$

Observe that if we introduce a new variable $y = x - l$,

$$\ddot{y} = -\frac{g}{l\sin\theta}y.$$

The above equation of motion describes a simple harmonic motion about an equilibrium position located at the bottom end of the ladder ($x = l$). The angular frequency of this simple harmonic motion is

$$\omega = \sqrt{\frac{g}{l\sin\theta}}.$$

The period is

$$T = 2\pi\sqrt{\frac{l\sin\theta}{g}}.$$

The time required for the ant to travel to the bottom end is a quarter of this period (as it begins at a point of greatest displacement with zero initial velocity and ends at the equilibrium position).

$$\frac{T}{4} = \frac{\pi}{2}\sqrt{\frac{l\sin\theta}{g}}.$$

11. A Spool**

Let the friction and the normal force between the spool and the ground be f and N respectively. Balancing torques about the center of the spool,

$$Fr = fR.$$

Balancing forces in the horizontal direction,

$$F\cos\theta = f$$

$$\implies \cos\theta = \frac{r}{R}$$

$$\theta = \cos^{-1}\left(\frac{r}{R}\right).$$

Balancing forces in the vertical direction,

$$N = mg - F\sin\theta.$$

At the maximum value of F, $f = \mu N$.

$$F\cos\theta = \mu(mg - F\sin\theta)$$

$$F = \frac{\mu mg}{\cos\theta + \mu\sin\theta} = \frac{\mu mgR}{r + \mu\sqrt{R^2 - r^2}}.$$

Given R, μ and r, the value of F that makes the spool slip is given by the expression above. To minimize this by varying r, the denominator needs to be maximised. The derivative of the denominator with respect to r is

$$1 - \frac{\mu r}{\sqrt{R^2 - r^2}} = 0.$$

Solving for the required value of r,

$$r = \frac{R}{\sqrt{1 + \mu^2}}.$$

12. Circle and Stick**

Let the friction force on the bottom of the circle be f rightwards. Then, the friction force between the circle and the stick must also be f for the torques on the circle to be balanced. Furthermore, in order for the horizontal forces on the combined system to be balanced, the friction on the stick due to the table must also be f leftwards. Let N_1, N_2 and N_3 be the normal forces on the circle due to the table, between the circle and the stick, and that on the stick by the table, respectively. Balancing torques on the stick about its bottom end,

$$\frac{3}{4}N_2 l = \frac{1}{2}mgl\cos 60^\circ$$

$$N_2 = \frac{1}{3}mg.$$

Balancing horizontal forces on the stick,

$$f(1+\cos 60^\circ) = N_2 \sin 60^\circ$$

$$f = \frac{\sqrt{3}}{9}mg.$$

Balancing vertical forces on the stick,

$$N_3 = mg - f\sin 60^\circ - N_2\cos 60^\circ$$

$$N_3 = \frac{2}{3}mg.$$

Balancing vertical forces on the circle,

$$N_1 = 3mg + N_2\cos 60^\circ + f\sin 60^\circ$$

$$N_1 = \frac{10}{3}mg.$$

Let μ_1, μ_2 and μ_3 be the coefficients of friction between the ground and the circle, between the circle and the stick, and between the stick and the ground, respectively. Then,

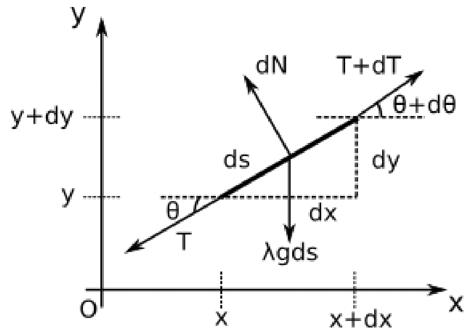
$$\mu_1 \geq \frac{f}{N_1} = \frac{\sqrt{3}}{30},$$

$$\mu_2 \geq \frac{f}{N_2} = \frac{\sqrt{3}}{3},$$

$$\mu_3 \geq \frac{f}{N_3} = \frac{\sqrt{3}}{6}.$$

13. Tension in a String**

Consider an infinitesimal segment of string between x-coordinates x and $x + dx$.



Balancing forces in the tangential direction (parallel to the segment),

$$(T + dT)\cos d\theta - T = dT = \lambda ds g \sin\theta.$$

Since $ds\sin\theta = \sqrt{1+y'^2}dx \cdot \frac{y'}{\sqrt{1+y'^2}} = y'dx$,

$$\int_{T_0}^{T} dT = \int_{x_0}^{x} \lambda g y' dx = \int_{y(x_0)}^{y} \lambda g dy$$

$$T = \lambda g \left(y(x) - y(x_0)\right) + T_0.$$

To apply the principle of virtual work, isolate the string segment between $x = x_0$ and $x = x$. Then, consider a virtual displacement δs of this string segment (towards $x = x$), tangential to the surface at all points. Then, the work done by the tensions on the ends of the strings is $(T-T_0)\delta s$. To compute the work done by gravity, observe that we have effectively transferred δs amount of string from a vertical level $y(x_0)$ to y. Thus, the change in potential energy is

$$\Delta U = \lambda g \left(y(x) - y(x_0)\right)\delta s.$$

By the principle of virtual work,

$$W_{tension} + W_{gravity} = 0$$

$$W_{tension} = -W_{gravity} = \Delta U$$

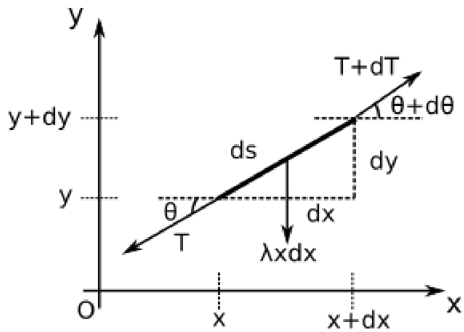
$$\implies T = \lambda g \left(y(x) - y(x_0)\right) + T_0.$$

$T(x)$ will remain the same after removing the surface. From the perspective of the first method, the tangential component of the forces will not change, resulting in the exact same $T(x)$. In the case of the second method, the string can still be displaced in an identical fashion to lead to the same work done by tension and gravity. Then, the same conclusion will naturally be reached. Interestingly, in the case of massless strings where $\lambda \to 0$,

$$T = T_0,$$

which shows that the tension is, in fact, uniform in a massless string that is resting on a surface or hanging under its own weight (this is intuitive as a massless string segment cannot experience a net force in the tangential direction).

14. Shape of String**



Consider the free-body diagram of a string segment between coordinates x and $x+dx$. Since there are no forces in the horizontal direction, the horizontal component of tension must be a constant F.

$$T\cos\theta = F.$$

Balancing forces in the vertical direction,

$$(T + dT)\sin(\theta + d\theta) - T\sin\theta = \lambda x dx.$$

Dividing both sides by dx,

$$\frac{d(T\sin\theta)}{dx} = \lambda x.$$

Since $T\sin\theta = F\tan\theta = Fy'$,

$$\frac{d\,(Fy')}{dx} = \lambda x$$

$$y'' = \frac{\lambda}{F}x$$

$$y' = \frac{\lambda}{2F}x^2 + y'(0)$$
$$y = \frac{\lambda}{6F}x^3 + y'(0)x + y(0).$$

15. Static Atwood**

Let the tension you exert be T, the tension exerted on the bottom of m_1 be T_1 and the tension on top be T_2. Intuitively, the minimum T occurs when m_2 tends to move downwards and m_1 tends to move upwards. Applying Eq. (7.3),

$$m_2 g \leq T_2 e^{\mu\theta}$$
$$T_1 \leq T e^{\mu\theta}.$$

Balancing the forces on m_1,

$$T_1 = T_2 - m_1 g \geq \frac{m_2 g}{e^{\mu\theta}} - m_1 g$$
$$\implies T \geq \frac{m_2 g}{e^{2\mu\theta}} - \frac{m_1 g}{e^{\mu\theta}}.$$

The above equality is satisfied when both of the first two inequalities reach their boundary case (impending motion) — this concurrency is definitely possible. Consider the case where the first equality is satisfied but not the second. At this juncture, T_2 has reached its minimum but T_1 has yet to reach its maximum. Then, it is possible for T_1 to increase further, thus increasing T_2 and preventing impending motion. A similar argument can be made in the reverse direction. Moving on, when T is maximum, the string tends to slip in the other direction. Then,

$$T_2 \leq m_2 g e^{\mu\theta}$$
$$T \leq T_1 e^{\mu\theta}$$
$$T_1 \leq m_2 g e^{\mu\theta} - m_1 g$$
$$T \leq m_2 g e^{2\mu\theta} - m_1 g e^{\mu\theta}.$$

An argument, similar to that above, can be made to show that the equality case is attainable. The maximum T is thus $m_2 g e^{2\mu\theta} - m_1 g e^{\mu\theta}$.

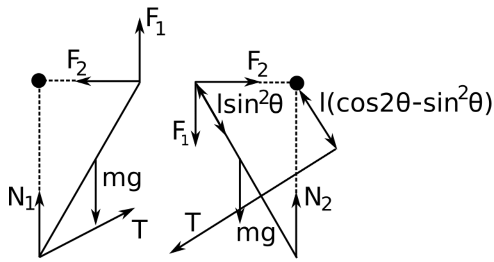


Figure 7.43: Three forces on the wedge

16. String and Sticks**

We divide the system into two sub-systems.

We take torques about the black circles depicted in Fig. 7.43 above, to eliminate the normal forces and F_2. Then,

$$F_1 \cdot l \sin\theta + T\cos\theta \cdot l\cos\theta = \frac{mgl\sin\theta}{2}$$

$$T(\cos 2\theta - \sin^2\theta)l = F_1 l \sin\theta + \frac{mg\sin\theta l}{2}.$$

Solving for T,

$$T = \frac{mg\sin\theta}{2\cos 2\theta}.$$

To solve this problem via the principle of virtual work, we use θ_0 to replace the angle θ defined in the problem and instead denote 2θ as a variable angle subtended by the two sticks. The idea here is to apply the principle of virtual work to the two sticks (excluding the string) after a virtual angular displacement $\delta\theta$ such that θ becomes $\theta + \delta\theta$. The virtual works done in this process are due to their weights (which can be computed as the negative of the total change in their gravitational potential energy) and the tensions acting on the two sticks. There is a slick way of computing the work done due to the latter factor — the work done by tension on the two sticks is simply the negative of the tension in the string, T, multiplied by the virtual extension of the string (negative as tension tugs on the sticks but an extension implies that the sticks move outwards, in the direction opposite to tension).

To this end, observe that the right end of the stick is located at a distance $l\cos 2\theta_0$ from the point of connection. Therefore, as a general function of θ, the length of the string is

$$s = \sqrt{l^2 + l^2\cos^2 2\theta_0 - 2l^2\cos 2\theta_0 \cos 2\theta}$$

by the cosine rule. The virtual extension due to the virtual angular displacement $\delta\theta$ is

$$\delta s = \frac{2l^2 \cos 2\theta_0 \sin 2\theta}{\sqrt{l^2 + l^2 \cos^2 2\theta_0 - 2l^2 \cos 2\theta_0 \cos 2\theta}} \delta\theta.$$

When $\theta = \theta_0$ initially, the virtual extension is

$$\delta s_0 = 2l \cos 2\theta_0 \delta\theta.$$

Moving on, the gravitational potential energy of the two sticks at angle θ, relative to the table, is

$$U = mgl \cos\theta.$$

The change in U due to an infinitesimal angular displacement $\delta\theta$ is

$$\delta U = -mgl \sin\theta\delta\theta = -W_G,$$

where W_G is the virtual work performed by gravity. When $\theta = \theta_0$ initially, the virtual work due to gravity is

$$W_G = mgl \sin\theta_0 \delta\theta.$$

Applying the principle of virtual work to the two sticks,

$$\begin{aligned} -T\delta s_0 + W_G &= 0 \\ T \cdot 2l \cos 2\theta_0 &= mgl \sin\theta_0 \\ T &= \frac{mg \sin\theta_0}{2\cos 2\theta_0}. \end{aligned}$$

17. Table**

The system has only one degree of freedom, θ, as the two legs must be symmetrical. A virtual displacement in θ results in there being no external non-conservative forces. Therefore, for the system to remain in static equilibrium, its total potential energy must be a stationary point. The total potential energy is due to the stretching of the spring and the gravitational potential energy of mass m.

$$\begin{aligned} U &= 2kl^2 \sin^2\theta + 2mgl \cos\theta, \\ \frac{dU}{d\theta} &= 4kl^2 \sin\theta \cos\theta - 2mgl \sin\theta. \end{aligned}$$

For the system to remain in static equilibrium,

$$\sin\theta(4kl^2\cos\theta - 2mgl) = 0.$$

$\theta = 0$ is a trivial equilibrium state. If $kl > \frac{mg}{2}$, another equilibrium position occurs when

$$\cos\theta = \frac{mg}{2kl}.$$

To examine the stability of these equilibriums, take the second derivative of U.

$$\frac{d^2U}{d\theta^2} = \cos\theta(4kl^2\cos\theta - 2mgl) - 4kl^2\sin^2\theta.$$

When $\theta = 0$,

$$\frac{d^2U}{d\theta^2} = 4kl^2 - 2mgl.$$

If $kl > \frac{mg}{2}$, the equilibrium is stable. Otherwise if $kl < \frac{mg}{2}$, the equilibrium is unstable. If $kl = \frac{mg}{2}$, one can show that all derivatives of U, evaluated at $\theta = 0$, are zero (as only the $\cos\theta$ terms matter when $\theta = 0$) — leading to a neutral equilibrium.

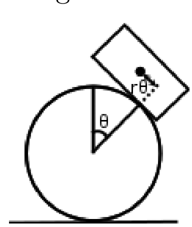
If $kl > \frac{mg}{2}$ such that another equilibrium position exists when $\cos\theta = \frac{mg}{2kl}$,

$$\frac{d^2U}{d\theta^2} = -4kl^2\sin^2\theta < 0.$$

Thus, the equilibrium is definitely unstable.

18. Block on Cylinder***

Consider the side-view corresponding to the cross-section of the cylinder. Let θ be the angle that the block has rotated. Then, θ is also the angle between the line joining the center of the cross-section to the point of contact and the vertical axis as shown in the figure.



Since the block does not slip relative to the cylinder, the point of contact must have been displaced by a distance $r\theta$ from the center of the relevant edge. To explicitly see why, the instantaneous center of rotation of the block is the point of contact (let its instantaneous coordinate be at θ). After the block rotates by an infinitesimal angle $d\theta$, the new point of contact is located at $\theta + d\theta$ with respect to the cylinder — thus the point of contact must have shifted by a distance $rd\theta$. Another perspective is that the cylinder "rolls" relative to the block (and we already know how to analyze this). Let the origin be located at the center of the circular cross section. Then, the y-coordinate of the center of mass of the block is

$$y = \left(r + \frac{h}{2}\right)\cos\theta + r\theta\sin\theta.$$

Since the gravitational potential energy of the block is directly proportional to y, we can simply compute the second derivative of y to check for stability (because static friction does no work).

$$\frac{dy}{d\theta} = -\left(r + \frac{h}{2}\right)\sin\theta + r\sin\theta + r\theta\cos\theta,$$

$$\frac{d^2y}{d\theta^2} = -\left(r + \frac{h}{2}\right)\cos\theta + 2r\cos\theta - r\theta\sin\theta.$$

At $\theta = 0$,

$$\frac{d^2y}{d\theta^2} = r - \frac{h}{2}.$$

Since $\frac{d^2y}{d\theta^2} > 0$ for a stable equilibrium,

$$r > \frac{h}{2}.$$

Chapter 8

Orbital Mechanics

This chapter will introduce Newton's law of gravitation and its ramifications for planetary motion — a phenomenon that was described by Kepler via his three empirical laws. Due to the inverse-squared nature of the law of gravitation, it has many aspects that are analogous to electrostatics. As such, we will mainly be discussing the central force problem and briefly run through the concepts of the gravitational field, potential and other relevant attributes of a mass distribution. The latter will be covered in-depth in the electrostatics chapter.

Before we embark on this topic proper, it is interesting to ponder why gravity — the weakest of the four fundamental forces — dominates interactions on the astronomical scale. The first factor is its infinite range — as opposed to the microscopic range of weak and strong interactions. However, this leaves yet another option — electromagnetic interactions — and leads us to the next factor. Gravitational interactions are dictated by mass, which can only be positive while charges in electromagnetism can be of opposite signs. The charge of a system on a galactic scale is rather neutral but the system is massive. Therefore, gravitational interactions outstrip electromagnetic interactions. That said, we can ask yet another question: why are large systems approximately neutral? Well, it is precisely because of the enormous strength of electromagnetic interactions that charges of opposite signs are pulled together — causing systems to be approximately neutral as a whole.

8.1 Newton's Law of Gravitation

Newton's law of gravitation states that every point mass attracts every other point mass. Explicitly, the gravitational force on a point mass m_1 by another point mass m_2 is

$$\boldsymbol{F}_{12} = -\frac{Gm_1m_2}{|\boldsymbol{r}_{21}|^2}\hat{\boldsymbol{r}}_{21} = -\frac{GMm}{|\boldsymbol{r}_{21}|^3}\boldsymbol{r}_{21}, \tag{8.1}$$

where $\boldsymbol{r}_{21}$ is the vector pointing from m_2 to m_1. $\boldsymbol{r}_{21} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the position vectors of m_1 and m_2 respectively. G is known as the universal gravitation constant and has a numerical value of $6.674 \times 10^{-11}\, m^3kg^{-1}s^{-2}$.

8.1.1 *Conserved Quantities in Planetary Motion*

Conservation of Energy

Now, consider the system of two masses m and M where $M \gg m$ such that M remains stationary at the origin (m could be a planet while M could be the Sun, for instance). We write the gravitational force on m as

$$\boldsymbol{F} = -\frac{GMm}{r^2}\hat{\boldsymbol{r}},$$

where $\boldsymbol{r}$ is the position vector of m. Observe that this takes the form of

$$\boldsymbol{F} = F(r)\hat{\boldsymbol{r}}.$$

That is, $\boldsymbol{F}$ is strictly in the radial direction and its magnitude is only dependent on the distance of m from the origin. Forces that can be expressed in such a form are known as central forces. They can be easily proven by the curl-test to be conservative. Thus, a potential energy can be associated with the gravitational force on m. If we imagine bringing m from infinity to its current radial distance r in a strictly radial manner (because the integral is path-independent) while maintaining M at the origin,

$$\begin{aligned} U &= -\int_{\infty}^{r} \boldsymbol{F} \cdot d\boldsymbol{r} \\ &= \int_{\infty}^{r} \frac{GMm}{r^2}\hat{\boldsymbol{r}} \cdot dr\hat{\boldsymbol{r}} \\ &= \left[-\frac{GMm}{r}\right]_{\infty}^{r}, \end{aligned}$$

where infinity has been defined to be the reference point at which the potential energy is zero. Thus,

$$U = -\frac{GMm}{r}. \tag{8.2}$$

Observe that the gravitational potential energy is a negative quantity. An intuitive explanation can be deduced from the fact that the potential energy is the work done by an external force in bringing mass m from infinity to its current state, without a change in kinetic energy. Since the external force must oppose the gravitational force which is attractive in nature, it must be directed radially outwards and thus performs negative work on m.

If m and M are the only interacting particles and if there are no external forces, the total mechanical energy E of m is conserved.

$$T + V = \frac{1}{2}mv^2 - \frac{GMm}{r} = E.$$

In light of the fact that the mass of the orbiting particle often cancels out in the end, we define the specific mechanical energy ε of m to be that obtained by dividing E by m. Furthermore, GM is defined to be the gravitational parameter μ.

$$\frac{1}{2}v^2 - \frac{\mu}{r} = \varepsilon.$$

ε is also conserved in m's resultant motion.

Conservation of Angular Momentum

Since the gravitational force on m is strictly radial, its angular momentum about the origin is conserved, such that

$$m\boldsymbol{r} \times \boldsymbol{v} = m\boldsymbol{r} \times \dot{\boldsymbol{r}} = \boldsymbol{L}$$

for some constant vector $\boldsymbol{L}$. Again, we define the specific angular momentum $\boldsymbol{h}$ to be the above expression divided by m, so that

$$\boldsymbol{r} \times \dot{\boldsymbol{r}} = \boldsymbol{h}$$

for another constant vector $\boldsymbol{h}$. From these conserved quantities alone, many problems can be solved.

Problem: The impact parameter b is defined to be the closest distance of approach between an incoming body and a central force source if the central force were to be switched off. If a body starts off at infinity, with an impact parameter b and initial velocity v_0, under the gravitational influence of a massive particle of mass M, determine the closest distance of approach r.

Let the velocity of the body at its closest approach be v. This must be perpendicular to its position vector with respect to M. By the conservation of angular momentum and energy,

$$bv_0 = vr,$$
$$\frac{1}{2}v_0^2 = \frac{1}{2}v^2 - \frac{\mu}{r}.$$

Substituting the expression for v in terms of r obtained from the first equation and simplifying,

$$v_0^2 r^2 + 2\mu r - b^2 v_0^2 = 0$$
$$r = \sqrt{\frac{\mu^2}{v_0^4} + b^2} - \frac{\mu}{v_0^2},$$

where the other negative root has been rejected, as it is physically incorrect.

8.2 Trajectory under Gravity

Before even manipulating any equations, we can actually greatly simplify a central force problem due to the conservation of angular momentum! This is because both the position vector $\boldsymbol{r}$ and velocity $\dot{\boldsymbol{r}}$ must be perpendicular to the fixed specific angular momentum vector $\boldsymbol{h}$ — that is, $\boldsymbol{r}$ and $\dot{\boldsymbol{r}}$ lie in a plane normal to $\boldsymbol{h}$. Since they have to lie in this plane at all instances, the central force problem can be reduced to a two-dimensional one.

Ideally, we wish to determine the trajectory of m in this plane in terms of Cartesian or polar coordinates. But before this, let us prove that its specific mechanical energy and angular momentum are conserved as consequences of its equation of motion. You might feel that this is trivial but this process is crucial later in extending our analysis to the case where both m and M can move, as the equation of motion takes a similar form. Applying Newton's second law to m and canceling the m's,

$$\ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r}.$$

Proof of Conserved Quantities Based on Equation of Motion

To prove the conservation of specific mechanical energy, take the dot product of $\dot{\boldsymbol{r}}$ with the equation of motion:

$$\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} = -\frac{\mu}{r^3}\boldsymbol{r} \cdot \dot{\boldsymbol{r}}.$$

To evaluate $\boldsymbol{r} \cdot \dot{\boldsymbol{r}}$, consider the following derivative.

$$\frac{d}{dt}(\boldsymbol{r} \cdot \boldsymbol{r}) = 2\boldsymbol{r} \cdot \dot{\boldsymbol{r}}$$
$$\implies \boldsymbol{r} \cdot \dot{\boldsymbol{r}} = \frac{1}{2}\frac{dr^2}{dt} = r\dot{r}.$$

Similarly,

$$\dot{\boldsymbol{r}} \cdot \ddot{\boldsymbol{r}} = \frac{1}{2}\frac{d}{dt}(\dot{\boldsymbol{r}} \cdot \dot{\boldsymbol{r}}) = \frac{1}{2}\frac{d(v^2)}{dt},$$

where v is the instantaneous speed of m. Note that v is different from $\dot{r}$ which is only the rate of change in the radial distance of m and does not take into account its tangential velocity. Substituting these expressions,

$$\frac{1}{2}\frac{d(v^2)}{dt} = -\frac{\mu}{r^2}\dot{r} = \frac{d}{dt}\left(\frac{\mu}{r}\right)$$
$$\frac{d}{dt}\left(\frac{1}{2}v^2 - \frac{\mu}{r}\right) = 0$$
$$\implies \frac{1}{2}v^2 - \frac{\mu}{r} = \varepsilon,$$

for some constant ε which we term the specific mechanical energy. To prove the conservation of angular momentum, consider the total time derivative of $\boldsymbol{r} \times \dot{\boldsymbol{r}}$.

$$\frac{d}{dt}(\boldsymbol{r} \times \dot{\boldsymbol{r}}) = \dot{\boldsymbol{r}} \times \dot{\boldsymbol{r}} + \boldsymbol{r} \times \ddot{\boldsymbol{r}} = \boldsymbol{0} + \boldsymbol{r} \times -\frac{\mu}{r^3}\boldsymbol{r} = \boldsymbol{0}$$
$$\implies \boldsymbol{r} \times \dot{\boldsymbol{r}} = \boldsymbol{h}$$

for some constant vector $\boldsymbol{h}$ which we call the specific angular momentum.

Trajectory

The trajectory of m can be derived via the following subtle manipulations — so follow closely. Firstly, take the cross-product of its equation of motion with $\boldsymbol{h}$.

$$\ddot{\boldsymbol{r}} \times \boldsymbol{h} = -\frac{\mu}{r^3}\boldsymbol{r} \times \boldsymbol{h}.$$

The left-hand side is the time derivative

$$\ddot{\boldsymbol{r}} \times \boldsymbol{h} = \frac{d}{dt}(\dot{\boldsymbol{r}} \times \boldsymbol{h}),$$

as $\boldsymbol{h}$ is constant. The right-hand side can also be expressed in terms of a total time derivative by using the BAC-CAB rule $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a} \cdot \boldsymbol{c}) - \boldsymbol{c}(\boldsymbol{a} \cdot \boldsymbol{b})$. Thus,

$$-\frac{\mu}{r^3}\boldsymbol{r} \times (\boldsymbol{r} \times \dot{\boldsymbol{r}}) = \frac{\mu}{r^3}\left[\dot{\boldsymbol{r}}(\boldsymbol{r} \cdot \boldsymbol{r}) - \boldsymbol{r}(\boldsymbol{r} \cdot \dot{\boldsymbol{r}})\right].$$

Substituting $\boldsymbol{r} \cdot \dot{\boldsymbol{r}} = r\dot{r}$,

$$\begin{aligned} -\frac{\mu}{r^3}\boldsymbol{r} \times (\boldsymbol{r} \times \dot{\boldsymbol{r}}) &= \frac{\mu}{r}\dot{\boldsymbol{r}} - \frac{\mu\dot{r}}{r^2}\boldsymbol{r} \\ &= \frac{d}{dt}\left(\frac{\mu}{r}\boldsymbol{r}\right). \end{aligned}$$

The original equation becomes

$$\begin{aligned} \frac{d}{dt}(\dot{\boldsymbol{r}} \times \boldsymbol{h}) &= \frac{d}{dt}\left(\frac{\mu}{r}\boldsymbol{r}\right) \\ \dot{\boldsymbol{r}} \times \boldsymbol{h} &= \frac{\mu}{r}\boldsymbol{r} + \mu\boldsymbol{e}, \end{aligned} \tag{8.3}$$

where $\mu\boldsymbol{e}$ has been set to be a constant vector of integration. $\boldsymbol{e}$ can be shown to lie in the plane of motion of m by taking the dot product of the equation above with $\boldsymbol{h}$. The first two expressions yield the null vector as they are perpendicular to $\boldsymbol{h}$ — implying that $\boldsymbol{e}$ follows suit. Finally, take the dot product of the above with $\boldsymbol{r}$.

$$\boldsymbol{r} \cdot (\dot{\boldsymbol{r}} \times \boldsymbol{h}) = \mu r + \mu\boldsymbol{r} \cdot \boldsymbol{e}.$$

The left-hand side can be simplified via the vector identity $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = \boldsymbol{b} \cdot (\boldsymbol{c} \times \boldsymbol{a}) = \boldsymbol{c} \cdot (\boldsymbol{a} \times \boldsymbol{b})$ such that

$$\boldsymbol{r} \cdot (\dot{\boldsymbol{r}} \times \boldsymbol{h}) = \boldsymbol{h} \cdot (\boldsymbol{r} \times \dot{\boldsymbol{r}}) = h^2.$$

Lastly, if θ is defined to be the angle that $\boldsymbol{r}$ makes with $\boldsymbol{e}$,

$$h^2 = \mu r + \mu r e \cos\theta.$$

Then,

$$r = \frac{\frac{h^2}{\mu}}{1 + e\cos\theta}. \tag{8.4}$$

Since $\boldsymbol{e}$ is defined to be a constant vector in the plane of motion of m, the x-axis of our fixed coordinate system can be defined to be along $\boldsymbol{e}$ such that Eq. (8.4) describes the trajectory of m in polar coordinates (r, θ). Finally, e

can in fact be expressed in terms of the conserved quantities in the previous section. Isolating $\boldsymbol{e}$ in Eq. (8.3),

$$\frac{1}{\mu}\dot{\boldsymbol{r}} \times \boldsymbol{h} - \frac{1}{r}\boldsymbol{r} = \boldsymbol{e}.$$

Taking the dot product of each side of the above with itself,

$$\left|\frac{1}{\mu}\dot{\boldsymbol{r}} \times \boldsymbol{h}\right|^2 + 1 - \frac{2}{\mu r}\boldsymbol{r} \cdot (\dot{\boldsymbol{r}} \times \boldsymbol{h}) = e^2.$$

Since $\dot{\boldsymbol{r}}$ is perpendicular to $\boldsymbol{h}$, the magnitude of $\dot{\boldsymbol{r}} \times \boldsymbol{h}$ is the product of their respective magnitudes vh where v is the instantaneous speed of m. The third term can be simplified into $\frac{2h^2}{\mu r}$ as we have shown that $\boldsymbol{r} \cdot (\dot{\boldsymbol{r}} \times \boldsymbol{h}) = h^2$. Thus,

$$\begin{aligned} e^2 &= \frac{v^2 h^2}{\mu^2} + 1 - \frac{2h^2}{\mu r} \\ &= 1 + \frac{2h^2}{\mu^2}\left(\frac{1}{2}v^2 - \frac{\mu}{r}\right) \\ &= 1 + \frac{2h^2 \varepsilon}{\mu^2}. \end{aligned}$$

Therefore, we have the following elegant result:

$$e = \sqrt{1 + \frac{2\varepsilon h^2}{\mu^2}}. \tag{8.5}$$

8.3 Conic Sections

Equation (8.4) in fact describes a conic section in general — a surface that is obtained from the intersection of a cone with a plane. In general, each conic section has one or two foci, points with respect to which the conic section has certain special properties. In polar coordinates, the equation of a general conic section with respect to one focus is

$$r = \frac{p}{1 + e\cos\theta}, \tag{8.6}$$

where θ is measured with respect to the x-axis. e determines the exact geometrical shape and is known as the eccentricity. p is half the length of a chord, parallel to the y-axis, that passes through the focus and is known as the semilatus rectum. The point that is closest to one of the foci is known as the periapsis and occurs when $\theta = 0$. The point, diametrically opposite to the periapsis and is furthest away from that focus, is known as the apoapsis and occurs when $\theta = \pi$ radians (if possible). Note that not all conic sections have an apoapsis as the furthest separation may tend to infinity (when $e \geq 1$).

We shall now show that the equations of various conic sections can be expressed in the form of Eq. (8.6) about a focus.

8.3.1 *Circles*

If $e = 0$,

$$r = p. \tag{8.7}$$

This is evidently the equation of a circle about the origin as the radial distance is independent of θ.

8.3.2 *Ellipses*

In the case where $0 < e < 1$, observe that the radius is no longer a constant but it still cannot tend to infinity (as the denominator is larger than zero). That is, the motion of m is still bounded to a certain region of space.

The equation of an ellipse whose center is defined at the origin is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1.$$

Without any loss of generality, suppose $a > b$ such that the ellipse is "fatter" than it is "taller." Then, a is known as the semi-major axis and is half the distance between the diametrically opposite points on the longer axis, which is the x-axis in this case. Similarly, b is known as the semi-minor axis and is half the distance between the diametrically opposite points on the shorter axis, which is the y-axis. Instead of a and b, an ellipse is often parameterized in terms of a and e, the latter of which was introduced as the eccentricity.

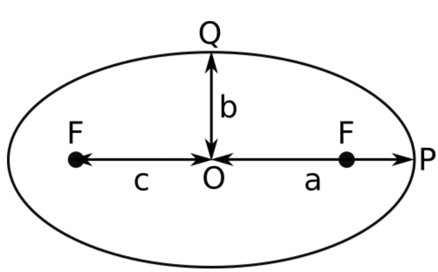


Figure 8.1: Ellipse

We define the eccentricity of an ellipse to be

$$e = \sqrt{1 - \frac{b^2}{a^2}},$$

such that

$$b = a\sqrt{1 - e^2}.$$

An ellipse also possesses the following unique property. Imagine tying the ends of an inextensible rope to two pins on a horizontal table (the rope is longer than the distance between the two pins). Then, the shape traced by picking up different segments of the rope and pulling it until it becomes taut is an ellipse. That is, the sum of the distances between every point on the ellipse and two points — known as the foci of the ellipse — is a constant. If the origin is defined at the center of the ellipse, the two foci lie on the longer axis and are symmetric about the shorter axis. The distance between a focus and the center is denoted as c. Then, we can compare the distances of a point Q along the y-axis and a point P along the x-axis to the respective focus to obtain c in terms of a and b.

$$2\sqrt{b^2 + c^2} = 2a$$
$$c^2 = a^2 - b^2.$$

In terms of e,

$$c = ae,$$

which is a pretty neat expression. Moving on to the main topic, we can show that Eq. (8.6) illustrates an ellipse about its right focus when $0 < e < 1$. If we define the origin to be at the right focus and (r, θ) to be polar coordinates about that focus,

$$x = r\cos\theta + c,$$
$$y = r\sin\theta,$$

where x and y are Cartesian coordinates with respect to the center of the ellipse. Substituting the above into the equation of an ellipse,

$$\frac{1}{a^2}(r^2\cos^2\theta + 2rc\cos\theta + c^2) + \frac{1}{b^2}r^2\sin^2\theta = 1.$$

Multiplying the above by b^2,

$$\frac{b^2}{a^2}(r^2\cos\theta^2 + 2rc\cos\theta + c^2) + r^2\sin^2\theta = b^2.$$

Using $b = a\sqrt{1-e^2}$ and $c = ae$,

$$(1-e^2)(r^2\cos^2\theta + 2era\cos\theta + a^2e^2) + r^2\sin^2\theta = a^2(1-e^2).$$

Rearranging,

$$r^2 = e^2r^2\cos^2\theta - 2(1-e^2)era\cos\theta + a(1-e^2)^2 = \left[er\cos\theta - a(1-e^2)\right]^2$$

$$r = \pm\left[er\cos\theta - a(1-e^2)\right].$$

Since r should be positive for the entire range of $0 \leq \theta \leq 2\pi$, we must take the negative expression (evident by substituting cases where $\cos\theta = 0$ or $\cos\theta = -1$).

$$r = a(1-e^2) - er\cos\theta$$

$$r = \frac{a(1-e^2)}{1+e\cos\theta}, \tag{8.8}$$

which is of the same form as Eq. (8.6).

8.3.3 *Parabola*

When $e = 1$, Eq. (8.6) becomes

$$r = \frac{p}{1+\cos\theta}.$$

This is in fact the equation of a parabola with the origin defined at its focus. Since r can tend towards infinity, the orbit of m is unbounded.

The equation of a parabola whose vertex is at the origin and is branching towards the left is

$$y^2 = -4ax.$$

a is in fact the distance between the vertex and the focus of the parabola, as we shall prove.

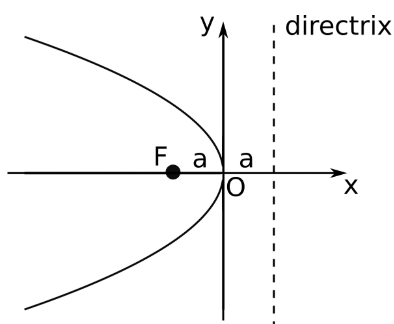


Figure 8.2: Parabola

A parabola is defined to be the set of points whose distances to a point, known as the focus, are identical to their perpendicular distances to a line known as the directrix. Let the focus be located at $(-a, 0)$ and the directrix be $x = a$ (Fig. 8.2). We will show that this results in a parabola illustrated by the equation above. Consider a point (x, y) on the parabola, $x < a$. Then, the above property implies that

$$\begin{aligned}\sqrt{(x+a)^2 + y^2} &= a - x \\ (x+a)^2 + y^2 &= x^2 - 2ax + a^2 \\ y^2 &= -4ax,\end{aligned}$$

which is the equation of a parabola with its vertex at the origin. We have shown that the distance between the focus and the vertex is a. The equation of a parabola with the origin defined at the focus is thus

$$y^2 = -4a(x - a).$$

Substituting $x = r\cos\theta$ and $y = r\sin\theta$ in polar coordinates,

$$\begin{aligned}r^2 \sin^2\theta &= -4a(r\cos\theta - a) \\ r^2 = r^2\cos^2\theta - 4ar\cos\theta + 4a^2 &= (r\cos\theta - 2a)^2 \\ r &= \pm(r\cos\theta - 2a).\end{aligned}$$

Since r should be positive when $\theta = \frac{\pi}{2}$, we must take the negative sign in the above expression.

$$\begin{aligned}r &= 2a - r\cos\theta \\ r &= \frac{2a}{1 + \cos\theta},\end{aligned} \tag{8.9}$$

which takes the form of Eq. (8.6) when $e = 1$.

8.3.4 *Hyperbola*

When $e > 1$, Eq. (8.6) yields

$$r = \frac{p}{1 + e\cos\theta}.$$

The radial distance of the particle m can again tend to infinity — implying that its orbit is, again, unbounded. As you might expect by now, the above equation describes half a hyperbola with the origin defined at its focus.

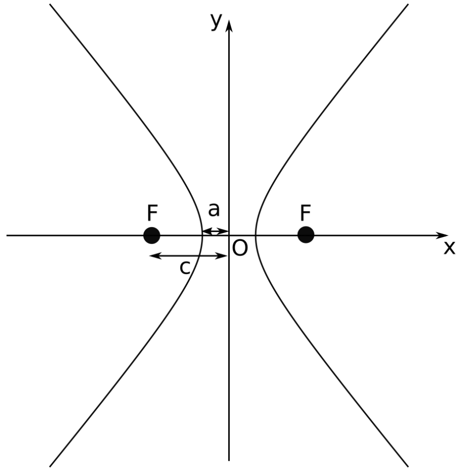


Figure 8.3: Hyperbola

The equation of a hyperbola whose vertex is located at the origin is

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

where a is the distance between the vertex and the periapsis. A hyperbola can also be defined in a different manner. The positive difference in the distances between every point on the hyperbola and its foci is a constant. The foci of a hyperbola are two points which lie on the line joining the periapsides and the vertex, and are symmetric about the vertex. Let them be located at $(c, 0)$ and $(-c, 0)$ respectively (Fig. 8.3). We will show that this property, along with the parameters c and a, will result in a hyperbola with its vertex at the origin. Consider a point along the x-axis that is supposed to be on the hyperbola (a periapsis). Evidently, this constant difference is

$$(c + a) - (c - a) = 2a.$$

Then, for a point (x, y) on the hyperbola in the region $x > 0$,

$$\sqrt{(x + c)^2 + y^2} - \sqrt{(x - c)^2 + y^2} = 2a.$$

Moving the second surd to the other side and squaring,

$$(x + c)^2 + y^2 = (x - c)^2 + y^2 + 4a^2 + 4a\sqrt{(x - c)^2 + y^2}$$
$$xc - a^2 = a\sqrt{(x - c)^2 + y^2}.$$

Squaring once again and simplifying yields

$$x^2c^2 + a^4 - 2a^2xc = a^2x^2 - 2a^2xc + a^2c^2 + a^2y^2$$
$$(c^2 - a^2)x^2 - a^2y^2 = a^2(c^2 - a^2).$$

Dividing the entire equation by $a^2(c^2 - a^2)$,

$$\frac{x^2}{a^2} - \frac{y^2}{c^2 - a^2} = 1.$$

Technically, we have only derived this for the region $x > 0$ but the negated version of our starting equation will result in the exact same expression for the region $x < 0$. We have shown that this definition of a hyperbola is valid. Comparing the denominators below the y^2 terms, it can be seen that

$$c^2 = a^2 + b^2.$$

Now if the eccentricity is defined as

$$e = \sqrt{1 + \frac{b^2}{a^2}}.$$

Then,

$$c = ae,$$
$$b = a\sqrt{e^2 - 1}.$$

We are now ready to show that Eq. (8.6) is hyperbolic, with the origin at its focus, if $e > 1$.

$$r = \frac{p}{1 + e\cos\theta};$$

r is evidently the minimum when $\theta = 0$. Hence, if the trajectory is really a hyperbola, it should be the half in the region $x < 0$. The equation of this half of the hyperbola about its focus is

$$\frac{(x - c)^2}{a^2} - \frac{y^2}{b^2} = 1.$$

Multiplying the above by b^2,

$$\frac{b^2}{a^2}(x^2 - 2xc + c^2) - y^2 = b^2.$$

Substituting the polar expressions for x and y, $c = ae$ and $b = a\sqrt{e^2-1}$,

$$(e^2-1)(r^2\cos^2\theta - 2r\cos\theta ae + a^2e^2) - r^2\sin^2\theta = a^2(e^2-1)$$

$$r^2 = e^2r^2\cos^2\theta - 2(e^2-1)r\cos\theta ae + a^2(e^2-1)^2 = \left[er\cos\theta - a(e^2-1)\right]^2$$

$$r = \pm\left[er\cos\theta - a(e^2-1)\right].$$

As r must be positive when $\theta = \frac{\pi}{2}$ radians, we take the negative expression in the above equation.

$$r = a(e^2-1) - er\cos\theta$$

$$r = \frac{a(e^2-1)}{1+e\cos\theta}. \tag{8.10}$$

8.3.5 *Dynamical Constants and Geometrical Properties*

We have shown that the shape of the orbit of a particle under a central gravitational force can only assume one of the four possible conic sections, excluding the trivial case of a straight line which occurs when $h = 0$. Now, we can relate the dynamical constants of motion — the specific mechanical energy ε and angular momentum h — to the shape of the orbit which is parameterized by the variables a (r in the case of a circle) and the eccentricity e.

Circular Orbits

Circular orbits occur when $e = 0$ — which occurs when ε and h satisfy a rather delicate relationship. Substituting the expression for e in Eq. (8.5),

$$1 + \frac{2h^2\varepsilon}{\mu^2} = 0$$

$$\varepsilon = -\frac{\mu^2}{2h^2}. \tag{8.11}$$

Furthermore, by comparing Eqs. (8.4) and (8.7),

$$r = \frac{h^2}{\mu},$$

$$h = \sqrt{\mu r}. \tag{8.12}$$

Substituting $h = \sqrt{\mu r}$ into Eq. (8.11),

$$\varepsilon = -\frac{\mu}{2r}, \tag{8.13}$$

which is intriguingly half the specific gravitational potential energy of m (i.e. its kinetic energy is half the negative value of its gravitational potential energy, $T = -\frac{1}{2}U$). Finally, note that these equations can be derived through much easier means by simply equating the gravitational force on m with its required centripetal force.

Elliptical Orbits

Elliptical orbits occur when $0 < e < 1$ which means that $\varepsilon < 0$. ε and h can be expressed in terms of a and e by comparing the corresponding variables in Eqs. (8.4) and (8.8), to get

$$\frac{h^2}{\mu} = a(1-e^2)$$
$$h = \sqrt{\mu a(1-e^2)}. \tag{8.14}$$

Applying Eq. (8.5),

$$1 - e^2 = -\frac{2h^2\varepsilon}{\mu^2}.$$

By Eq. (8.14), the left-hand side is $\frac{h^2}{\mu a}$. Thus,

$$-\frac{2h^2\varepsilon}{\mu^2} = \frac{h^2}{\mu a}$$
$$\varepsilon = -\frac{\mu}{2a}. \tag{8.15}$$

The specific mechanical energy of an elliptical orbit is so important that a name has been coined for its equivalent form. Applying the conservation of energy,

$$\frac{1}{2}v^2 - \frac{\mu}{r} = -\frac{\mu}{2a}$$
$$v^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right). \tag{8.16}$$

This is known as the vis-viva equation which relates the speed of particle m to its radial position in an elliptical orbit.

Parabolic Orbits

The orbit of a particle is a parabola when $e = 1$ and thus, occurs when

$$\varepsilon = 0. \tag{8.17}$$

Intuitively, the particle possesses sufficient total mechanical energy to overcome the potential energy barrier (the maximum potential energy is zero and occurs at infinity). Comparing Eqs. (8.4) and (8.9),

$$\frac{h^2}{\mu} = 2a$$
$$h = \sqrt{2a\mu}. \tag{8.18}$$

Hyperbolic Orbits

A hyperbolic orbit arises when $e > 1$ and thus when $\varepsilon > 0$. Comparing Eqs. (8.4) and (8.10),

$$\frac{h^2}{\mu} = a(e^2 - 1)$$
$$h = \sqrt{\mu a(e^2 - 1)}. \tag{8.19}$$

Substituting the expression for e in Eq. (8.5) in the first equation,

$$\frac{h^2}{\mu} = a\frac{2h^2\varepsilon}{\mu^2}$$
$$\varepsilon = \frac{\mu}{2a}. \tag{8.20}$$

8.4 Kepler's Three Laws

Following from the above derivations, it is easy to prove Kepler's three empirical laws which describe the motion of planets about the Sun. In this case, we assume the mass of the Sun M to be much larger than that of a single planet, such that to a good approximation, the net gravitational force on each planet is only due to that exerted by the Sun.

First Law: Planets travel in elliptical orbits around the Sun as a focus.

Proof: Since the orbits of planets must be bounded, they must generally travel in an ellipse around a focus (where M is located) as shown in the previous section.

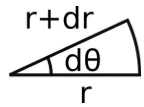


Figure 8.4: Infinitesimal sector

Second Law: The vector pointing from the Sun to a planet sweeps out areas at a constant rate as the planet orbits around the Sun.

Proof: This is just an embellished statement of the conservation of angular momentum of the planet about the Sun. Let the instantaneous radial distance between the planet and the Sun be r. Then, consider an infinitesimal area dA swept by the radial vector when the planet undergoes an infinitesimal angular displacement $d\theta$. dA is the area of a sector of radius r and angle $d\theta$ (Fig. 8.4).

$$dA = \frac{1}{2}r^2 d\theta. \tag{8.21}$$

Dividing both sides by dt,

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} = \frac{1}{2}r^2\dot{\theta}.$$

Since $h = r^2\dot{\theta}$,

$$\frac{dA}{dt} = \frac{h}{2}. \tag{8.22}$$

Thus, the rate of area swept is a constant.

Third Law: The square of the period of the orbit of a planet, T, is proportional to the cube of the semi-major axis length, a. Concretely,

$$T^2 = \frac{4\pi^2 a^3}{\mu}. \tag{8.23}$$

Proof: Separating variables in Eq. (8.22) and integrating over the entire ellipse and one period,

$$\int_0^{\pi ab} dA = \int_0^T \frac{h}{2}dt$$

$$\pi ab = \frac{h}{2}T,$$

where we have used the fact that the area of an ellipse with a semi-major axis length a and semi-minor axis length b is πab. Squaring both sides,

$$\pi^2 a^2 b^2 = \frac{h^2}{4} T^2.$$

Substituting $b^2 = a^2(1-e^2)$ and $\frac{h^2}{\mu} = a(1-e^2)$ for an ellipse,

$$\pi^2 a^4 (1-e^2) = \frac{1}{4}\mu a (1-e^2) T^2$$
$$T^2 = \frac{4\pi^2 a^3}{\mu}.$$

8.5 Two-Body Problem

Having analyzed the central gravitational force problem where $m \ll M$, we return to the more general problem where M also moves under the gravitational influence of m. This is known as the two-body problem. Writing the equations of motion for both masses while using $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ to denote the position vectors of m and M respectively,

$$m\ddot{\boldsymbol{r}}_1 = -\frac{GMm}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3}(\boldsymbol{r}_1 - \boldsymbol{r}_2),$$
$$M\ddot{\boldsymbol{r}}_2 = -\frac{GMm}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|^3}(\boldsymbol{r}_2 - \boldsymbol{r}_1).$$

Adding the two equations together,

$$m\ddot{\boldsymbol{r}}_1 + M\ddot{\boldsymbol{r}}_2 = (m+M)\ddot{\boldsymbol{r}}_{\boldsymbol{CM}} = 0.$$

As expected, the center of mass of the two particles travels at a constant velocity. Since the center of mass of the particles lies along the line joining them, the particles must orbit about their common center of mass with the **same angular velocity**. Next, consider the center of mass frame which is an inertial frame. We still use $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ to represent the position vectors of m and M in this frame for the sake of convenience. In the center of mass frame,

$$m\boldsymbol{r}_1 = -M\boldsymbol{r}_2,$$

where the origin is set at the center of mass, so to solve for the entire system, we can just determine how $\boldsymbol{r}_1 - \boldsymbol{r}_2$ evolves with time. We shall let this separation vector be $\boldsymbol{r}$. Multiplying the first equation of motion by M and subtracting it by the second, which is multiplied by m,

$$mM\ddot{\boldsymbol{r}} = -\frac{GMm(M+m)}{r^3}\boldsymbol{r}$$
$$\frac{mM}{m+M}\ddot{\boldsymbol{r}} = -\frac{GMm}{r^3}\boldsymbol{r}. \tag{8.24}$$

Observe that this differential equation is of the same form as that of a body with mass m orbiting a fixed mass M (such that $\mu = GM$) given by

$$m\ddot{\boldsymbol{r}} = -\frac{\mu m}{r^3}\boldsymbol{r},$$

after substituting $m = \frac{mM}{m+M}$ and $\mu = G(m+M)$. We essentially have an effective mass $m_{eq} = \frac{mM}{m+M}$, known as the reduced mass, orbiting a mass $m+M$ fixed at the origin; $\boldsymbol{r}$ then denotes the position vector of this reduced mass m_{eq} in the center of mass frame. Therefore, the results from the above section are directly applicable and the trajectory of $\boldsymbol{r}$ in polar coordinates in the center of mass frame is

$$r = \frac{\frac{h^2}{\mu}}{1+e\cos\theta},$$

where $\boldsymbol{h} = \boldsymbol{r}\times\dot{\boldsymbol{r}}$ is the specific angular momentum of the equivalent mass m_{eq} in the center of mass frame. The conservation of h is ensured as the equation of motion takes a similar form as before — implying that the motion of the two bodies are again confined to a single plane. In fact, you can prove that $\boldsymbol{h}$ is the total angular momentum of m and M in the center of mass frame divided by m_{eq}. Next, e is analogously

$$e = \sqrt{1+\frac{2\varepsilon h^2}{\mu^2}},$$

with

$$\varepsilon = \frac{1}{2}v^2 - \frac{\mu}{r},$$

where $v = |\dot{\boldsymbol{r}}|$ is the instantaneous speed of m_{eq} in the center of mass frame. ε is indeed a constant as a result of the equation of motion. In fact, it can be proven that ε is the total mechanical energy of the original system of m and M in the center of mass frame, divided by the reduced mass m_{eq}.

Moving on, since $m\boldsymbol{r}_1 = -M\boldsymbol{r}_2$ such that $\boldsymbol{r}_1 = \frac{M}{m+M}\boldsymbol{r}$ and $\boldsymbol{r}_2 = -\frac{m}{m+M}\boldsymbol{r}$, the trajectories of m and M in the center of mass frame are

$$r_1 = \frac{\frac{Mh^2}{(m+M)\mu}}{1+e\cos\theta},$$

$$r_2 = \frac{\frac{mh^2}{(m+M)\mu}}{1-e\cos\theta},$$

with the same h and e. The negative sign in front of e in the second expression stems from the fact that the position vector $\boldsymbol{r}_2$ of M is anti-parallel to $\boldsymbol{r}$. Actually, it is a flipped version of the position vector of m about the common center of mass (after scaling by $\frac{m}{M}$). This reflection also implies that the common center of mass sits on different foci with respect to the individual elliptical trajectories of m and M (if their orbits are bounded — a phenomenon that is very likely). That is, the common center of mass will lie at the right focus of one ellipse and at the left focus of the other.

Finally, the orbits of the two bodies are **only unbounded if their total mechanical energy is at least zero in the center of mass frame** because ε, which dictates whether their orbits are confined (as it determines the magnitude of e), is directly proportional to the total mechanical energy of the combined system in the center of mass frame (it was remarked that it is in fact the total mechanical energy divided by m_{eq}).

The Lagrangian Perspective

Actually, the above manipulations become more enlightening in the Lagrangian formulation (see Chapter 12). For two particles m and M whose interaction potential energy is of the form $U(\boldsymbol{r}_1 - \boldsymbol{r}_2)$ where $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ are the position vectors of m and M, the Lagrangian of the system comprising the two particles is

$$\mathcal{L} = \frac{1}{2}m\dot{\boldsymbol{r}}_1^2 + \frac{1}{2}M\dot{\boldsymbol{r}}_2^2 - U(\boldsymbol{r}_1 - \boldsymbol{r}_2).$$

Now, introduce the new variables $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$ and $\boldsymbol{R} = \frac{m\boldsymbol{r}_1 + M\boldsymbol{r}_2}{m+M}$, where the latter is the position vector of the center of mass. Then, $\boldsymbol{r}_1 = \boldsymbol{R} + \frac{M}{m+M}\boldsymbol{r}$ and $\boldsymbol{r}_2 = \boldsymbol{R} - \frac{m}{m+M}\boldsymbol{r}$ such that

$$\mathcal{L} = \frac{1}{2}m\left(\dot{\boldsymbol{R}} + \frac{M}{m+M}\dot{\boldsymbol{r}}\right)^2 + \frac{1}{2}M\left(\dot{\boldsymbol{R}} - \frac{m}{m+M}\dot{\boldsymbol{r}}\right)^2 - U(\boldsymbol{r})$$

$$\mathcal{L} = \frac{1}{2}(m+M)\dot{\boldsymbol{R}}^2 + \frac{1}{2}\frac{mM}{m+M}\dot{\boldsymbol{r}}^2 - U(\boldsymbol{r}),$$

as the cross-terms cancel out. We have effectively decoupled the Lagrangian into two parts which are solely functions of $\boldsymbol{R}$ and $\boldsymbol{r}$, respectively. Observing that the components of $\boldsymbol{R}$ are cyclic coordinates (i.e. there is nothing in the Lagrangian that is directly a function of the components of $\boldsymbol{R}$), we can conclude that $\dot{\boldsymbol{R}}$ is a constant (i.e. the center of mass moves with a constant velocity). Refer to Chapter 12 for more details about this. Meanwhile, the other portion of the Lagrangian that is a function of $\boldsymbol{r}$ suggests that the system can be treated as a single particle with a reduced mass $m_{eq} = \frac{mM}{m+M}$ whose position vector is $\boldsymbol{r}$ relative to the origin, under the influence of a central potential energy given by $U(\boldsymbol{r})$. We have thus reduced a two-body central force problem into a one-body problem! In the particular case of $U(\boldsymbol{r}) = -\frac{GMm}{r}$, the above implies that the reduced mass is effectively interacting with a fixed mass $(m + M)$ at the origin (as $-\frac{G \cdot \frac{mM}{m+M} \cdot (m+M)}{r} = -\frac{GMm}{r}$)!

8.5.1 *Modified Kepler's Laws*

Having drawn an analogy between the central force problem and the two-body problem, it is easy to prove the following modified set of Kepler's laws for an isolated system of masses m and M in the frame of their center of mass.

First Law: m and M travel in elliptical orbits about their common center of mass as a focus.

Second Law: The vector pointing from the common center of mass of m and M to either m or M sweeps out area at a constant rate.

Third Law: The square of T, the common period of the orbits of m and M, is proportional to the cube of the semi-major axis length of m or M (denoted as a_1 or a_2). Concretely,

$$T^2 = \frac{4\pi^2 a_{eq}^3}{\mu} = \frac{4\pi^2 (m+M)^2 a_1^3}{M^3} = \frac{4\pi^2 (m+M)^2 a_2^3}{m^3}, \tag{8.25}$$

where $a_{eq} = \frac{m+M}{M} a_1 = \frac{m+M}{m} a_2$ is the semi-major axis of the ellipse traversed by the reduced mass and $\mu = G(m + M)$ is the equivalent gravitational parameter.

8.6 Mass Distributions

The previous section only analyzed the gravitational interaction between two discrete particles. In the more general case, we would like to consider the effect of an entire mass distribution. The crucial component in connecting the effects of individual masses is the principle of superposition. It states that the gravitational force on a mass due to a mass distribution is the sum of the gravitational forces due to each individual mass component on that particular mass. The total gravitational force on a mass M due to N discrete masses ranging from m_1 to m_N is

$$\boldsymbol{F} = -\sum_{i=1}^{N} \frac{GMm_i}{r_i^2}\hat{\boldsymbol{r}}_{\boldsymbol{i}},$$

where $\boldsymbol{r_i}$ is a vector pointing from m_i to M. For a continuous distribution,

$$\boldsymbol{F} = -\int \frac{GM}{r^2}\hat{\boldsymbol{r}}dm,$$

where $\boldsymbol{r}$ is the vector pointing from each infinitesimal mass element dm to M. The integral is performed over the entire mass distribution. In light of the above expression, a vector field $\boldsymbol{g}$, known as the gravitational field, can be defined for all points in space to compute the force per unit mass of a particle placed at any point in space, due to a predetermined mass distribution.

$$\boldsymbol{g} = -\sum_{i=1}^{N} \frac{Gm_i}{r_i^2}\hat{\boldsymbol{r}}_{\boldsymbol{i}} = -\int \frac{G}{r^2}\hat{\boldsymbol{r}}dm, \tag{8.26}$$

where $\boldsymbol{r_i}$ and $\boldsymbol{r}$ point from each mass component to the point of concern. The force on a mass M placed at a point where the gravitational field is $\boldsymbol{g}$ due to the other masses, is then

$$\boldsymbol{F} = M\boldsymbol{g}.$$

The infinitesimal gravitational flux through an infinitesimal area $d\boldsymbol{A}$ is the dot product of $\boldsymbol{g}$, the gravitational field strength at that point and the infinitesimal area vector.

$$d\Phi = \boldsymbol{g} \cdot d\boldsymbol{A}.$$

The direction of an area vector is arbitrary for an open surface. However for closed surfaces, the area vector is defined to be directed outwards by convention. The total gravitational flux cutting through a closed surface S is then

$$\Phi = \oint\!\!\!\oint_S \boldsymbol{g} \cdot d\boldsymbol{A},$$

where the loop around the integral indicates that S is a closed surface. Due to the inverse-squared nature of $\boldsymbol{g}$, the total flux emanating through any closed surface is in fact directly proportional to the amount of mass enclosed inside the closed surface, M_{enc}. This is known as Gauss's law of gravitation which states that

$$\Phi = \oiint_S \boldsymbol{g} \cdot d\boldsymbol{A} = -4\pi G M_{enc}. \tag{8.27}$$

Gauss's law provides an elegant way of computing $\boldsymbol{g}$ for symmetric mass distributions. Another way of computing $\boldsymbol{g}$ which is convenient for non-symmetric distributions is derived from the conservative nature of $\boldsymbol{g}$. The gravitational potential at a position $\boldsymbol{r}$ due to a mass distribution is defined as

$$V(\boldsymbol{r}) = -\int_{\infty}^{\boldsymbol{r}} \boldsymbol{g} \cdot d\boldsymbol{r}, \tag{8.28}$$

where infinity has been taken as a reference point at which the potential is zero. The above provides a way of calculating V, given $\boldsymbol{g}$. However, V can be determined directly from the mass distribution by noting that each individual mass component dm contributes

$$-\int_{\infty}^{\boldsymbol{r}} -\frac{Gdm}{r^2}\hat{\boldsymbol{r}} \cdot d\boldsymbol{r} = \int_{\infty}^{r} \frac{Gdm}{r^2} dr = -\frac{Gdm}{r}$$

to the potential at a distance r away from the mass component. Thus, the total potential at a point due to a mass distribution can also be written as

$$V = \int -\frac{G}{r} dm \tag{8.29}$$

by the principle of superposition, where r is the distance between an infinitesimal mass element dm and the point of concern. For discrete distributions involving N masses, the potential at a particular point is

$$V = \sum_{i=1}^{N} -\frac{Gm_i}{r_i}, \tag{8.30}$$

where r_i is the distance between m_i and the point of concern. Supposing that we are able to calculate V for all points in space via the above process, we can conversely compute $\boldsymbol{g}$ by taking the negative gradient of V, as a result of Eq. (8.28). In Cartesian, cylindrical and spherical coordinates, the negative gradients of V are, respectively,

$$\boldsymbol{g} = -\frac{\partial V}{\partial x}\hat{\boldsymbol{i}} - \frac{\partial V}{\partial y}\hat{\boldsymbol{j}} - \frac{\partial V}{\partial z}\hat{\boldsymbol{k}},$$

$$\boldsymbol{g} = -\frac{\partial V}{\partial r}\hat{\boldsymbol{r}} - \frac{1}{r}\frac{\partial V}{\partial \theta}\hat{\boldsymbol{\theta}} - \frac{\partial V}{\partial z}\hat{\boldsymbol{k}},$$

$$\boldsymbol{g} = -\frac{\partial V}{\partial r}\hat{\boldsymbol{r}} - \frac{1}{r\sin\theta}\frac{\partial V}{\partial \phi}\hat{\boldsymbol{\phi}} - \frac{1}{r}\frac{\partial V}{\partial \theta}\hat{\boldsymbol{\theta}}.$$

Next, by observing Eq. (8.28), one can notice that the potential at a point in space due to a fixed mass distribution is the potential energy per unit mass of a mass that is placed at that point, due to the fixed mass distribution. Thus, the potential energy of a mass M due to a fixed mass distribution at a point with potential V is

$$U = -\int_{\infty}^{\boldsymbol{r}} M\boldsymbol{g} \cdot d\boldsymbol{r} = MV. \tag{8.31}$$

The force on mass M due to the fixed mass distribution is similarly the negative gradient of U. Finally, the total potential energy of a mass distribution is, by definition, the sum of the gravitational potential energy between each pair of particles (with no double-counting of pairs). In the discrete case of N charged particles, let U_{ij} be the potential energy associated with the interactions between particle i and j.

$$U_{ij} = -\frac{Gm_i m_j}{r_{ij}},$$

where r_{ij} is the distance between the particles. The total potential energy of the distribution is then

$$U = \sum_{i,j\, i<j} -\frac{Gm_i m_j}{r_{ij}}, \tag{8.32}$$

where the $i < j$ below the summation sign prevents double-counting. We can in fact double-count each pair and include the factor of half such that

$$U = \frac{1}{2}\sum_{i=1}^{N} m_i \sum_{j\neq i} -\frac{Gm_j}{r_{ij}} = \frac{1}{2}\sum_{i=1}^{N} m_i V_i, \tag{8.33}$$

where V_i is the potential at the position of the ith particle, due to the other particles. In the case of certain continuous mass distributions, the contribution due to an infinitesimal mass element to the potential at its own position is negligible. Then, we can include its own contribution in applying the formula above such that the total potential energy due to a mass distribution is

$$U = \frac{1}{2} \int V dm. \tag{8.34}$$

That is, we can sum the product of the potential with the infinitesimal mass at each point in space, over the entire distribution.

The above is a brief overview of the concepts involving a mass distribution. For a more detailed explanation, refer to the chapter on electrostatics.

Problems

Orbits

1. *Vis-viva Equation**

Prove the vis-viva equation (Eq. (8.16)), via the conservation of energy and angular momentum.

2. *Planet**

A planet is currently at the perihelion of its orbit around the Sun (the point at which its distance to the Sun is shortest). The perihelion distance from the Sun is R_p and the velocity of the planet at this point is v_p. Find the velocity of the planet at the aphelion (the point where its distance to the Sun is largest) through the vis-viva equation.

3. *Impact Parameter**

A body starts off at infinity with an impact parameter b and initial velocity v_0 under the gravitational influence of a massive particle of mass M. Its orbit is evidently a hyperbola as its total mechanical energy is positive. Show that the impact parameter b is also the b in the hyperbola equation that describes the orbit of the body. Determine the eccentricity of the orbit and the angle of deflection of the body (the angle between its initial and final velocities).

4. *Circular Orbit**

The trajectory of a particle under the influence of a central force (not necessarily gravitational) is a circle of radius R. The position of the source of the central force is not known. Given that the maximum and minimum speeds of the particle are v_1 and v_2 respectively, determine the period T of the particle's orbit.

5. *Binary Stars**

Two binary stars are moving in circular orbits around their common center of mass, with a period T. If the two stars are somehow stopped by an external agency at a certain juncture and then gently released, determine the time of collision τ between the two stars after their release.

6. *Hohmann Transfer**

The Earth is currently orbiting around the Sun, which has mass M, at a radius r_1. A Hohmann transfer is used to switch an orbiting body from one circular orbit to another. It is performed by two instantaneous tangential boosts (instantaneous changes in tangential velocity) with the first boost setting the orbiting body into an elliptical path and the second boost, at the apogee of the ellipse, returning the orbiting body into the required circular orbit. Supposing that we wish to launch a satellite from the North Pole of the Earth to a circular orbit of radius $r_2 > 2r_1$ around the Sun, determine the time required for the transfer and the sum of the two instantaneous changes in speed. Ignore any gravitational effects on the satellite due to the Earth. Qualitatively propose a way to reduce the sum of these speed boosts.

7. *Missile**

Consider the system of a satellite of mass m_1 orbiting around the Earth of mass $M \gg m_1$ and radius R_e. In the frame of the Earth, the satellite is orbiting in an ellipse with semi-major axis length a and eccentricity e. A missile of mass $m_2 \ll M$ is launched from the surface of the Earth, travels along a straight line and sticks to the satellite at the apogee (furthest point from the Earth in the elliptical orbit). If the satellite then moves off to infinity, determine the minimum initial velocity of the missile in the frame of the Earth, u. Neglect any gravitational effects between m_1 and m_2 and of the Sun.

8. *Tilting Ellipse**

A body is orbiting elliptically about a massive body of mass M, with a semi-major axis length a_0 and eccentricity e_0. Suppose that when the body is at the apoapsis, it is given a sudden boost such that its radial velocity is instantly increased to u. If the resultant orbit is still an ellipse, determine its semi-major axis length a_1, eccentricity e_1 and the acute angle between the final and initial semi-major axes.

Mass Distributions

9. *Extracting a Mass*

A point mass m is currently stationary at an arbitrary position inside a uniform, spherical shell of mass M and radius r. Determine the minimum work performed by an external force in extracting the point mass m through a small hole on the shell and bringing it to infinity.

10. ***Ring and Mass****

A stationary point mass m is placed a distance h above the center of a stationary, uniform ring of mass M and radius r, along the symmetrical axis. Determine the velocity of m when it reaches the center of the ring.

11. ***Tunnel Through Earth*****

Imagine that you dug a hole from one side of the earth to the opposite side (not necessarily through the center of mass). The path makes a perpendicular distance h with the center of the Earth. You then jump into the hole with a negligible initial velocity. Assuming that you survive the process and that the Earth is a sphere with a constant mass density ρ and radius R, prove that you will undergo simple harmonic motion. Determine the angular frequency of this oscillation.

12. ***Atmosphere*****

Model the Earth's atmosphere as an isothermal and uniform ideal gas of temperature T and suppose that its thickness is comparable to the radius of the Earth. Show that the equilibrium pressure of the atmosphere at an altitude h above the surface of the Earth obeys

$$p = p_0 e^{-\frac{GM_0 Mh}{RTR_e(R_e+h)}},$$

where M_0 is the average molar mass of the molecules in the Earth's atmosphere, R is the ideal gas constant, R_e is the radius of the Earth and p_0 is the pressure at the surface of the Earth. Assume that the Earth is a uniform sphere and neglect any gravitational effects due to the atmosphere and the rotation of the Earth. Hint: the ideal gas law states that $pV = nRT$ where p is the pressure, V is the volume of the gas and n is the number of moles.

Solutions

1. Vis-viva Equation*

Let the distances of the periapsis and apoapsis to the focus at which the massive body lies be r_p and r_a respectively. Let the velocity of the orbiting body at these points be u and v respectively. By the conservation of angular momentum and energy,

$$ur_p = vr_a,$$

$$\frac{1}{2}u^2 - \frac{\mu}{r_p} = \frac{1}{2}v^2 - \frac{\mu}{r_a}.$$

Substituting the expression for v obtained from the first equation into the second and solving for u,

$$\frac{1}{2}u^2 = \frac{\mu r_a}{(r_a + r_p)r_p} = \frac{\mu r_a}{2ar_p},$$

since $r_a + r_p = 2a$. Substituting this into the equation for the specific mechanical energy of the orbiting body at the periapsis,

$$\varepsilon = \frac{1}{2}u^2 - \frac{\mu}{r_p} = \frac{\mu(r_a - 2a)}{2ar_p} = \frac{-\mu r_p}{2ar_p} = -\frac{\mu}{2a}.$$

Therefore, for an arbitrary radial distance r, the speed of the orbiting body is given by

$$v^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right).$$

2. Planet*

Let the semi-major axis length be a. From the vis-viva equation,

$$\frac{1}{2}v_p^2 - \frac{\mu}{r_p} = -\frac{\mu}{2a}$$

$$a = \frac{\mu r_p}{2\mu - v_p^2 r_p}.$$

The distance between the Sun and the aphelion is

$$r_a = 2a - r_p = \frac{v_p^2 r_p^2}{2\mu - v_p^2 r_p}.$$

By the conservation of angular momentum,

$$r_a v_a = r_p v_p.$$

Thus, the velocity at the aphelion is

$$v_a = \frac{r_p v_p}{r_a} = \frac{2\mu}{v_p r_p} - v_p.$$

3. Impact Parameter*

Recall that the distance of closest approach is

$$r = \sqrt{\frac{\mu^2}{v_0^4} + b^2} - \frac{\mu}{v_0^2}.$$

This quantity is $c - a$ where c is the distance between the focus and the vertex and a is the distance between the vertex and the periapsis.

$$c - a = \sqrt{\frac{\mu^2}{v_0^4} + b^2} - \frac{\mu}{v_0^2}.$$

Furthermore, we know from the relationship between the geometry of the hyperbola and the dynamical constants that

$$\varepsilon = \frac{1}{2}v_0^2 = \frac{\mu}{2a}$$
$$\implies \frac{\mu}{v_0^2} = a$$
$$\implies c - a = \sqrt{a^2 + b^2} - a$$
$$\implies c^2 = a^2 + b^2.$$

Let the value of b in the equation of the hyperbola be b'. From the definition of c,

$$c^2 = a^2 + b'^2.$$

Therefore,

$$b = b'.$$

From the geometry of the orbit,

$$b'^2 = a^2(e^2 - 1)$$
$$e = \sqrt{\frac{b'^2}{a^2} + 1} = \sqrt{\frac{b^2 v_0^4}{\mu^2} + 1}.$$

Let the x-axis be normal to the hyperbola at its periapsis. Then, the angle that the initial velocity of the body makes with the x-axis is

$$\tan^{-1}\frac{b}{a} = \tan^{-1}\frac{bv_0^2}{\mu}.$$

Since the hyperbola is symmetrical about the x-axis, the angle of deflection is π minus twice of the angle above.

$$\theta_{def} = \pi - 2\tan^{-1}\frac{bv_0^2}{\mu}.$$

4. Circular Orbit*

Firstly, observe that the central force source must be positioned within the circle. Suppose that this is otherwise in search of a contradiction. Then, one can draw a line through the central force source that intersects the circle at two distinct points. The velocities of the particle at these points result in different directions of angular momenta about the source (this must be the case so that the orbit of the particle is a continuous circle) — contradicting the fact that the angular momentum of a particle under the sole influence of a central force is conserved about the central force.

Now that we know that the central force source must be located within the circle, define the origin at the source and orient the x and y-axes such that the center of the circle is at coordinates $(a, 0)$ with $0 < a < R$. By the conservation of specific angular momentum h about the source,

$$v_1(R - a) = v_2(R + a) = h.$$

Solving for h,

$$h = \frac{2Rv_1v_2}{v_1 + v_2}.$$

The rate of area swept by the position vector of the particle is

$$\frac{dA}{dt} = \frac{h}{2} = \frac{Rv_1v_2}{v_1 + v_2}$$

by Eq. (8.22). Separating variables and integrating over a single period T,

$$\int_0^{\pi R^2} dA = \int_0^T \frac{Rv_1v_2}{v_1 + v_2} dt$$

$$T = \frac{\pi R(v_1 + v_2)}{v_1v_2},$$

where we have used the fact that the total area swept by the position vector of the particle in a single period is the area bounded by its circular trajectory, πR^2, since the central force source is within the circle.

5. Binary Stars*

Let the distance between the two binary stars be r and their masses be m and M. Applying the notion of a reduced mass, the period T of this binary star system is akin to a reduced mass $m_{eq} = \frac{mM}{m+M}$ orbiting around a fixed mass $m+M$, located at the common center of mass, in a circle of radius r. By Kepler's third law,

$$T^2 = \frac{4\pi^2 r^3}{G(m+M)}.$$

Next, when the two stars are stopped and gently released, they travel in a straight line towards each other. Introducing a reduced mass again, the time of collision τ is equivalent to the time that m_{eq} takes to collide with the fixed mass $m+M$, located at the common center of mass, given that it starts at zero initial velocity and initial distance r. This can be considered as half the period of a flattened elliptical orbit of m_{eq} with eccentricity $e \to 1$, such that a focus is located at the periapsis (where the fixed mass $m+M$ is), and semi-major axis length $\frac{r}{2}$. By Kepler's third law,

$$(2\tau)^2 = \frac{4\pi^2 \left(\frac{r}{2}\right)^3}{G(m+M)}.$$

Dividing this by the previous equation and taking square roots on both sides,

$$\tau = \frac{T}{4\sqrt{2}}.$$

6. Hohmann Transfer**

The semi-major axis length of the ellipse is

$$a = \frac{r_1 + r_2}{2}.$$

By Kepler's third law, the period of the elliptical orbit is

$$T = \sqrt{\frac{\pi^2 (r_1 + r_2)^3}{2\mu}}.$$

The time taken for the mission is half the period

$$\frac{T}{2} = \pi\sqrt{\frac{(r_1 + r_2)^3}{8\mu}}.$$

Let the velocity of the satellite after the first boost be u and the velocity of the satellite at the apogee be v. By the conservation of angular momentum and energy,

$$ur_1 = vr_2,$$

$$\frac{1}{2}u^2 - \frac{\mu}{r_1} = \frac{1}{2}v^2 - \frac{\mu}{r_2},$$

$$\frac{1}{2}\left(1 - \frac{r_1^2}{r_2^2}\right)u^2 = \frac{\mu}{r_1} - \frac{\mu}{r_2}.$$

Thus,

$$u = \sqrt{\frac{2\mu r_2}{(r_1 + r_2)r_1}}.$$

The initial velocity u_0 of the satellite (before the first boost) in a circular orbit essentially of radius r_1 can be obtained by considering the centripetal force required for its orbit.

$$\frac{u_0^2}{r_1} = \frac{\mu}{r_1^2}$$

$$u_0 = \sqrt{\frac{\mu}{r_1}}.$$

Thus the first boost results in an increase in speed of magnitude

$$u - u_0 = \sqrt{\frac{2\mu r_2}{(r_1 + r_2)r_1}} - \sqrt{\frac{\mu}{r_1}}.$$

The speed of the satellite at the apogee is

$$v = \frac{r_1}{r_2}u = \sqrt{\frac{2\mu r_1}{(r_1 + r_2)r_2}}.$$

The required speed of circular motion at radius r_2 is $v' = \sqrt{\frac{\mu}{r_2}}$. Thus, the second boost is

$$v' - v = \sqrt{\frac{\mu}{r_2}} - \sqrt{\frac{2\mu r_1}{(r_1 + r_2)r_2}}.$$

The sum of the increases in speed is thus

$$(r_2 - r_1)\sqrt{\frac{2\mu}{(r_1 + r_2)r_1 r_2}} + \sqrt{\frac{\mu}{r_1 r_2}}(\sqrt{r_1} - \sqrt{r_2}).$$

To optimize the transfer, one can conduct the launch near the equator at an opportune time such that the rotation of the Earth increases the initial velocity of the satellite (u_0). To ensure that the satellite's velocity is purely tangential with respect to the Sun, one can perform the boost only when the launch site is at the point on the equator that is closest to or furthest away from the Sun (in determining these locations and the correct times, it is essential to note that Earth's axis of rotation is currently tilted by 23.5° with respect to its axis of revolution around the Sun).

7. Missile**

Let v_r be the velocity of the missile at the apogee, right before the collision. Note that this is solely in the radial direction. By the conservation of energy with $\mu = GM$,

$$\frac{1}{2}u^2 - \frac{\mu}{R_e} = \frac{1}{2}v_r^2 - \frac{\mu}{a(1+e)},$$

as the apogee distance is $a(1+e)$ by Eq. (8.8).

$$v_r^2 = u^2 + \frac{2\mu}{a(1+e)} - \frac{2\mu}{R_e}.$$

Now, let the velocity of the satellite at the apogee be v_t. By the vis-viva equation,

$$v_t^2 = \mu\left(\frac{2}{a(1+e)} - \frac{1}{a}\right).$$

The radial and tangential velocities of the satellite-cum-missile immediately after the collision are

$$v_r' = \frac{m_2}{m_1 + m_2}v_r,$$
$$v_t' = \frac{m_1}{m_1 + m_2}v_t$$

by the conservation of momentum. Its total post-collision energy is

$$\frac{1}{2}(m_1+m_2)v_r'^2+\frac{1}{2}(m_1+m_2)v_t'^2-\frac{\mu(m_1+m_2)}{a(1+e)}$$
$$=\frac{m_2^2}{2(m_1+m_2)}u^2-\frac{\mu}{2(m_1+m_2)}\left(\frac{m_1^2}{a}+\frac{2m_2^2}{R_e}+\frac{4m_1m_2}{a(1+e)}\right).$$

The total mechanical energy must be larger than zero for the system to travel to infinity. Hence, the minimum u is

$$u=\sqrt{\frac{\mu}{m_2^2}\left(\frac{m_1^2}{a}+\frac{2m_2^2}{R_e}+\frac{4m_1m_2}{a(1+e)}\right)}.$$

8. Tilting Ellipse**

Since the specific angular momentum h remains the same,

$$\frac{h^2}{\mu}=a_0(1-e_0^2)=a_1(1-e_1^2).$$

Furthermore, relating the specific mechanical energy to the semi-major axes length of the orbits,

$$\frac{1}{2}v^2-\frac{\mu}{r_a}=-\frac{\mu}{2a_0}$$
$$\frac{1}{2}(u^2+v^2)-\frac{\mu}{r_a}=-\frac{\mu}{2a_1},$$

where $r_a=a_0(1+e_0)$ is the distance between the focus and the apoapsis. Subtracting the above equations,

$$\frac{\mu}{2}\left(\frac{1}{a_0}-\frac{1}{a_1}\right)=\frac{1}{2}u^2.$$

Solving,

$$a_1=\frac{\mu a_0}{\mu-a_0u^2}.$$

Substituting this expression for a_1 into the first equation,

$$e_1=\sqrt{1-\frac{(\mu-a_0u^2)(1-e_0^2)}{\mu}}.$$

To determine the angle that the semi-major axis has rotated, consider the new trajectory equation.

$$r=\frac{a_1(1-e_1^2)}{1+e_1\cos\theta}=\frac{a_0(1-e_0^2)}{1+e_1\cos\theta}.$$

Define θ_0 to be the angle between the new semi-major axis and the original apoapsis. Then,

$$a_0(1+e_0) = \frac{a_0(1-e_0^2)}{1+e_1\cos\theta_0},$$

where $a_0(1+e_0)$ is the original apoapsis distance.

$$\cos\theta_0 = -\frac{e_0}{e_1},$$

θ_0 is the obtuse angle between the semi-major axes. Thus, the acute angle between them is

$$\theta = \cos^{-1}\frac{e_0}{e_1} = \cos^{-1}\frac{e_0}{\sqrt{1-\frac{(\mu-a_0u^2)(1-e_0^2)}{\mu}}}.$$

9. Extracting a Mass*

By drawing a concentric spherical Gaussian surface within the spherical shell, one can conclude that the gravitational field within the shell, due to the shell is zero. Thus, the gravitational potential due to the shell is equal for all points within the shell. Consequently, the gravitational potential energy associated with the interactions between the shell and mass m can be computed as that of a mass m at the center of the shell. Since all points on the shell are equidistant from the center, the potential energy and thus the total mechanical energy of the combined system is

$$E = -\frac{GMm}{r}.$$

The minimum amount of work is performed when m is stationary at infinity. At this juncture, the mechanical energy of the combined system is zero. Hence, the minimum work performed by an external force is

$$W = 0 - E = \frac{GMm}{r}.$$

10. Ring and Mass**

Since the distances between all points on the ring and the point mass are equal, the initial mechanical energy of the combined system is

$$E = -\frac{GMm}{\sqrt{r^2+h^2}}.$$

When m reaches the center of the ring, let the velocities of m and M be v_1 and v_2 respectively. Since the velocity of the center of mass is constant due to the lack of external forces and was originally zero,

$$mv_1 + Mv_2 = 0.$$

Finally, the total mechanical energy of the system must be conserved, hence

$$\frac{1}{2}mv_1^2 + \frac{1}{2}Mv_2^2 - \frac{GMm}{r} = -\frac{GMm}{\sqrt{r^2+h^2}}.$$

Solving this set of equations,

$$v_1 = \sqrt{\frac{2GM^2\left(\frac{1}{r} - \frac{1}{\sqrt{r^2+h^2}}\right)}{m+M}},$$

along the direction pointing from its initial position to the center of the ring.

11. Tunnel Through Earth**

Consider a side-view of the situation — the tunnel is a chord in a circle of radius r. Let x denote the distance along this chord, where $x = 0$ has been defined to be at the center of the chord. When you are at a coordinate x, your distance from the center of the sphere is $r = \sqrt{h^2 + x^2}$. Drawing a Gaussian surface corresponding to a spherical shell of this radius, the gravitational field strength at this point can be computed as

$$g \cdot 4\pi r^2 = -4\pi G\rho\frac{4}{3}\pi r^3$$
$$g = -\frac{4\pi}{3}G\rho r.$$

The force on you in the direction of the chord is

$$mg \cdot \frac{x}{r} = -\frac{4\pi G\rho m}{3}x.$$

Therefore, your equation of motion is

$$\ddot{x} = -\frac{4\pi G\rho}{3}x,$$

which represents a simple harmonic motion. The angular frequency is

$$\omega = \sqrt{\frac{4\pi G\rho}{3}}.$$

12. Atmosphere**

Let $\rho(r)$ be the density of the atmosphere at a distance r from the center of the Earth. Consider the forces on an infinitesimal volume element at a distance r from the center of the Earth, in spherical coordinates. The difference in pressure multiplied by the cross-sectional area of the element must balance the gravitational force on it. Then,

$$-\frac{G\rho M}{r^2}dV = dpdA$$

$$\frac{dp}{dr} = -\frac{G\rho M}{r^2}.$$

Note that ρ is also a function of r which can be expressed in terms of p via the ideal gas law

$$pV = nRT.$$

Rearranging,

$$p = \frac{\rho}{M_0}RT.$$

Hence,

$$\frac{dp}{dr} = -\frac{GM_0M}{RTr^2}p$$

$$\int_{p_0}^{p} \frac{1}{p}dp = \int_{R_e}^{R_e+h} -\frac{GM_0M}{RTr^2}dr.$$

Thus,

$$p = p_0 e^{-\frac{GM_0Mh}{RTR_e(R_e+h)}}.$$

Chapter 9

Fluids

This chapter will analyze the statics and dynamics of fluids — a form of matter that has yet to be considered. Due to the immense complexity of this field, our analysis in fluid dynamics will only be limited to the special case of steady, incompressible and energy-conserving flow.

9.1 Properties of a Fluid

The defining feature of a fluid is its vulnerability to deformations. An ideal fluid with no viscosity deforms under the influence of forces[1] parallel to its surface — known as shear forces — until it takes on a certain form such that there are no longer any shear forces on it. This occurs both when the fluid is static and when it is flowing. Therefore, forces on a steady-state fluid must be perpendicular to its surface.

Since a conservative force is the negative gradient of its associated potential energy, the surface of a fluid at equilibrium under the sole influence of conservative forces must be equipotential. Therefore, if we neglect the effects of viscosity and surface tension of real fluids and model the Earth as a uniform sphere (assumed to be non-rotating as well), the surface of a beaker of water should really be a spherical cap if we zoom in close enough.

9.1.1 *Pressure*

Instead of forces, it is convenient to define the pressure P due to a fluid on a surface. The pressure on a surface due to a force is the magnitude of the

[1]Note that this does not refer to a net force aggregated over all surfaces but the sheer existence of a net force at a single point on a surface.

normal component of force per unit area on it, i.e.

$$P = \frac{dF}{dA}. \tag{9.1}$$

We shall now understand why pressure is defined as a scalar, instead of a vector.

Isotropy of Pressure

The pressure at each point in a fluid is isotropic (i.e. identical in all directions). Consider an infinitesimal wedge from its side view in the xz-plane (Fig. 9.1). The cross-section of the wedge is uniform along the y-direction.

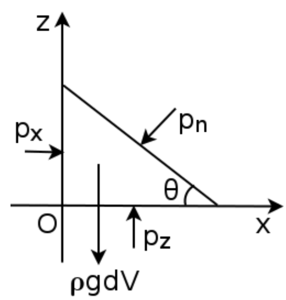


Figure 9.1: Side-view of infinitesimal wedge

Let the lengths of the edges of the wedge along the corresponding axes be dx, dy and dz respectively. Note that even though pressure only tells us about the normal component of force at each surface, this problem is still completely determined by the fact that the fluid cannot withstand any shear forces (so there are no shear forces). Applying Newton's second law to the z-direction while considering the forces due to pressure and the weight of the element,

$$p_z dxdy - p_n \cdot \frac{dx}{\cos\theta} dy \cos\theta - \rho dV g = \rho dV a,$$

where ρ is the density of the fluid and a is the acceleration of the fluid element in the z direction. dV is the volume of the infinitesimal volume element, $dV = \frac{1}{2} dxdydz$. Taking the limits $dx \to 0$, $dy \to 0$ and $dz \to 0$,

$$p_z = p_n.$$

Applying a similar process to the x direction would yield

$$p_x = p_n.$$

Therefore,

$$p_x = p_z = p_n = p,$$

for some constant p. Since we did not define a particular value of θ, we can vary it while ensuring that the wedge encloses a point of concern, to conclude that the pressure along a neighbourhood in the xz-plane about that point is some constant p. Next, to prove that the pressure is p in all planes in the neighborhood of a particular point, consider another infinitesimal fluid element as shown below.

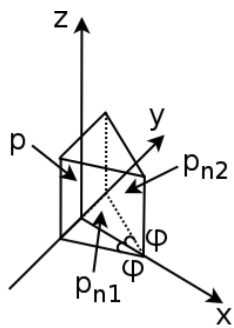


Figure 9.2: Infinitesimal wedge

One can easily show that $p_{n1} = p_{n2} = p$. Then, by varying ϕ while enclosing a point of concern, we can conclude that the pressure is p along all planes in the vicinity of that point and is hence isotropic. Due to its isotropy, pressure is defined to be a scalar as a particular direction does not need to be assigned to it.

9.2 Fluid Statics

As usual, if a fluid is in static equilibrium, there must be no net external force on each element of the fluid. In a stationary container where the only forces on each fluid element are its weight, the normal force due to the container and the force due to its surrounding pressure, the pressure varies with depth throughout the fluid. We first show that the pressure at all points of the same horizontal level must be identical. Consider an infinitesimal cube element with edge lengths dx, dy and dz (Fig. 9.3).

Balancing the forces on the cube along the x and y directions, we conclude that the pressure must be constant throughout the entire xy-plane, as long as the fluid is still continuous. Now, consider the forces on the cube in the vertical direction.

$$dp \cdot dA = -\rho dAdzg$$
$$\frac{dp}{dz} = -\rho g.$$

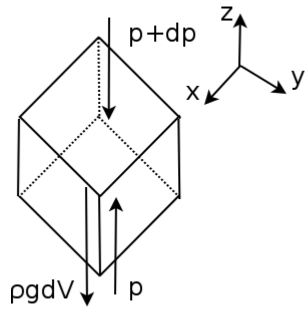


Figure 9.3: Infinitesimal cube

If ρ is uniform,

$$p = p_{ref} - \rho g z, \tag{9.2}$$

for a continuous vertical column of fluid. p_{ref} is the pressure at coordinates $z = 0$. This equation is also applicable to moving fluids whose elements are not accelerating vertically. The combination of the previous two properties implies that the pressure of a continuous fluid in static equilibrium is always $p_{ref} - \rho g z$, regardless of how oddly shaped the container may be.

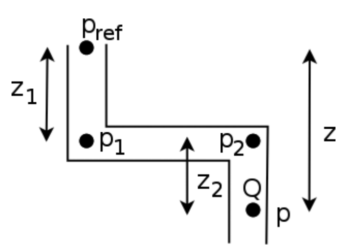


Figure 9.4: Fluid in container

Consider the container in the figure above. The pressure at the top is p_{ref} and we wish to determine the pressure p at point Q. We know that

$$p_1 = p_{ref} + \rho g z_1,$$

$$p = p_2 + \rho g z_2$$

from the previous result. Furthermore, we know that

$$p_1 = p_2$$

as these points are at the same vertical level. Thus,

$$p = p_{ref} + \rho g(z_1 + z_2) = p_{ref} + \rho g z.$$

9.2.1 *Archimedes' Principle*

Due to the variation in pressure, a body partially or fully submerged in a stationary fluid experiences an upward force called the upthrust or buoyant force. Archimedes' Principle states that a body experiences an upthrust equal to the weight of the fluid it displaces. Note that this upthrust acts at the center of gravity[2] of the displaced fluid (also known as the center of buoyancy) and is directed vertically upwards.

$$F_{upthrust} = m_{dis}g = \rho_{fluid}V_{dis}g. \tag{9.3}$$

An intuitive and arguably, most insightful, proof of Archimedes' Principle is that the fluid surrounding the object does not "know" whether the fluid, that was originally present, has been removed or is still there. Thus, the remaining fluid will exert the same force as before on whatever entity occupies the space where part of the fluid originally was. For the displaced fluid to remain at rest previously, the upthrust must have balanced its weight and passed through its center of gravity (for torque balance). Thus, the submerged object experiences a force upwards whose magnitude is commensurate with the weight of the displaced fluid, acting on the center of gravity of the displaced fluid. Note that this argument works for general gravitational fields since we did not assume the gravitational field to be uniform.

Problem: A stationary cylinder of mass m and base area A is partially submerged in a fluid of uniform density ρ, with its cylindrical axis aligned with the vertical. Supposing that the cylinder is given a slight vertical displacement, determine the angular frequency of small oscillations.

Let x be the length of the cylinder that is currently submerged in the fluid. Then the equation of motion of the cylinder is

$$m\ddot{x} = mg - \rho Axg.$$

Using the substitution $u = x - \frac{m}{\rho A}$,

$$\ddot{u} = -\frac{\rho Ag}{m}u.$$

Thus, the angular frequency is

$$\omega = \sqrt{\frac{\rho Ag}{m}}.$$

[2] The center of gravity of an extended body is the point at which its weight appears to act at. It is generally different from the center of mass but is equal to the center of mass when the body is placed in a uniform gravitational field.

9.2.2 *Moving Containers*

When a container is undergoing a certain form of motion, there must be no relative motion between the fluid it contains and itself at equilibrium (and also between adjacent fluid elements). This is due to the viscosity of the fluid in realistic situations which acts to oppose the relative motion between a fluid element and a neighboring surface, in a manner analogous to friction. In light of this necessity for the fluid to move in tandem with the container, the variation in fluid pressure must be consistent with the motion of the fluid. This also affects the shape of the surface of the fluid as pressure is proportional to the depth of a fluid. Consider the rotating container below as an example.

Problem: Calculate the height h of the equilibrium liquid level in a cylinder of liquid, rotating at a constant angular velocity ω, as a function of radial distance r from the axis of rotation, which is the cylindrical axis. The incompressible liquid has a uniform density ρ, and h at $r = 0$ is h_0.

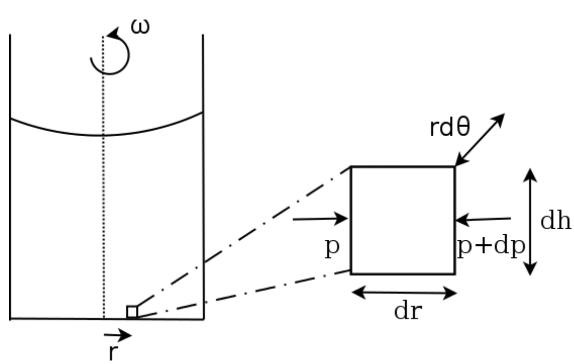


Figure 9.5: Rotating fluid

We consider an infinitesimal element of liquid with mass $dm = \rho r dr d\theta dh$ at the bottom of the cylinder, in cylindrical coordinates. We note that the net force due to the differences in pressure must provide the centripetal force in order for this element to rotate at angular velocity ω. Considering the forces in the radial direction,

$$p r dh d\theta - (p + dp) r dh d\theta = -dm r\omega^2 = -\rho r^2 \omega^2 dr d\theta dh$$

$$\frac{dp}{dr} = \rho r \omega^2.$$

Since $p = \rho g h$,

$$\frac{dh}{dr} = \frac{r\omega^2}{g}$$

$$\int_{h_0}^{h} dh = \int_0^r \frac{r\omega^2}{g} dr$$

$$h = h_0 + \frac{r^2\omega^2}{2g},$$

where h_0 is the height of the fluid level along the cylindrical axis. It is interesting to note that this means that the surface of the water is a paraboloid. Actually, this problem can be easily solved by introducing a notion known as the centrifugal potential and imposing the condition that the water surface must be equipotential at equilibrium (see Chapter 11 Problem 8).

9.3 Surface Tension

Surface tension results from the discrepancy in the cohesive interactions between molecules of a liquid near its interface with another medium and between those within the liquid. Consider a liquid that is coexisting with its vapor state[3] while assuming that the interactions between the liquid and vapor phases are negligible.

In a homogeneous liquid, the generally attractive[4] interactions between a particular molecule and all other molecules ascribe a negative potential energy to that particular molecule — known as the binding energy. The negative of the binding energy is then the external work needed to remove a molecule from the liquid to infinity. It happens that the potential energy associated with a pair of molecules decreases rapidly with their internuclear distance such that the potential energy of a single molecule is largely due to its immediate neighbours. With these observations, we can create a simple microscopic depiction of the origin of a concept known as surface energy.

Referring to Fig. 9.6, a liquid molecule near the interface has roughly half as many neighboring liquid molecules to interact with, as compared to one

[3]Note that even if we isolated a liquid in vacuum, part of it will always evaporate to form a thin vapor film over the liquid surface.

[4]Note that the interactions are highly repulsive in the regime of short internuclear distances such that, although two isolated molecules will actually be attracted from far away, when they get close enough, there will be a certain separation where the net force between them is zero.

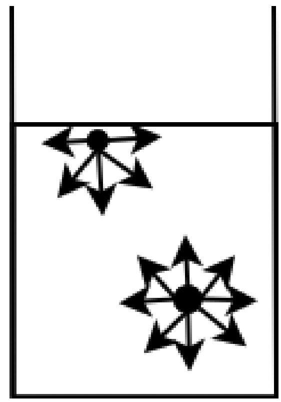


Figure 9.6: Molecules in a liquid

in the interior of the liquid. Then, if the binding energy of a molecule within the liquid is E (which is a negative value), the binding energy of a surface molecule is only $\frac{E}{2}$. This "missing" negative energy ($\frac{E}{2}$ per surface molecule) can then be supplemented to the surface of a liquid as an additional positive potential energy, known as the surface energy. Moreover, this model also suggests that the surface energy should be proportional to the area of the surface, as a larger surface naturally contains more surface molecules.

In light of the above discussion, we can define a quantity known as the surface tension γ to describe the surface energy U stored per unit area of the surface of a liquid.

$$\gamma = \frac{dU}{dA}. \tag{9.4}$$

At an interface between two homogeneous fluids, the surface tension does not depend on how much a surface has already been stretched. Since γ is constant, the surface energy of a liquid can be calculated as

$$U = \gamma A. \tag{9.5}$$

As the surface energy is positive, a liquid surface will naturally attempt to reduce its surface area to minimize its total potential energy (while taking into account other factors such as gravitational potential energy). Next, another equivalent perspective is that the surface tension is the external work required in increasing the area of the liquid surface by a unit area by moving molecules from the interior (where they have a more negative binding energy) to the surface (where they have a less negative binding energy).

$$\gamma = \frac{dW_{ext}}{dA}. \tag{9.6}$$

Now, where there is energy, we would expect there to be a force as well. To this end of developing a force perspective of surface tension, consider the following instructive set-up.

A liquid film is surrounded by a "U-shaped" frame and one of its ends is pierced by a thin metal wire of length l that is pulled by a force F, such that the wire moves at a constant velocity. Notice that since the wire does not accelerate, F must balance the force due to surface tension on the wire — implying that we only have to determine F to find the force associated with surface tension.

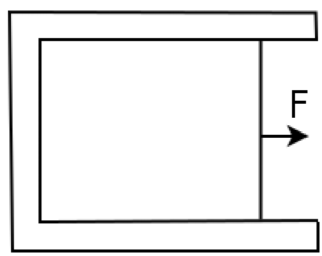


Figure 9.7: Liquid film

To solve for F, we can apply the principle of virtual work[5] to the wire. Suppose that the wire is displaced rightwards by a virtual displacement δx. The work done by the external force on the wire would be $F\delta x$ while the work done by the liquid film would be the negative change in surface energy, $-2\gamma dA = -2\gamma l\delta x$. The factor of two arises from the fact that there are two liquid-vapor interfaces — one on top of the liquid and one below. Since the wire is in equilibrium, the principle of virtual work asserts that the sum of virtual works yields zero so that

$$F\delta x - 2\gamma l\delta x = 0$$
$$\implies F = 2\gamma l.$$

Therefore, the liquid film tugs at the wire with a force $2\gamma l$ leftwards. Since there are two liquid-air interfaces, we can ascribe a surface tension force of γl to each interface. Now, observe that if we scale the length of the wire by a factor k, the surface tension force due to each interface will

[5] Alternatively, we can apply Eq. (9.6).

also be scaled by k. Therefore, a horizontal strip of liquid with vertical length dl would exert a force γdl on the wire. Since each interface of the original set-up can be viewed as the composition of myriad horizontal strips of liquid that are identical, and whose total exerted force is γl, the meaning of γ must be the force per unit length of a line along a liquid surface!

A paramount observation at this juncture is that γ must not only be the force per unit length at the end near the wire but it must also be the force per unit length for all lines along the entire liquid interface. That is, a form of internal tension must exist on the entire liquid surface to ensure that all sections remain at equilibrium (similar to how an internal tension $T = F$ must exist within a rigid rod when its ends are pulled by an external force F).

As hinted by the previous set-up, we propose the following alternative view of surface tension. For an arbitrary liquid-vapor interface S, draw a line L along it which divides it into two regions (Fig. 9.8).

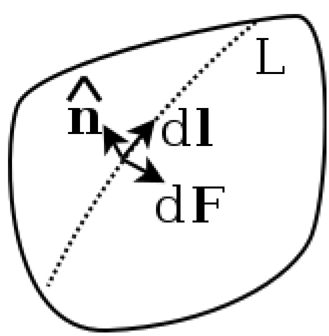


Figure 9.8: Force due to infinitesimal line segment

The contribution to the force between the regions due to an infinitesimal line segment $\boldsymbol{dl}$ along the interface is

$$d\boldsymbol{F} = \gamma d\boldsymbol{l} \times \hat{\boldsymbol{n}},$$

where $\hat{\boldsymbol{n}}$ is the normal vector of the interface at that line segment. The above is the force that the region, corresponding to the direction of $d\boldsymbol{F}$, exerts on the other region along the segment. For example, since $d\boldsymbol{F}$ points towards the right for the infinitesimal segment depicted above, it is the force exerted by the right region on the left region along the segment $d\boldsymbol{l}$. Note that the direction of $\hat{\boldsymbol{n}}$ is arbitrary, but it is a good habit to define a consistent direction (such as outwards). Finally, it is crucial to observe that $d\boldsymbol{F}$ is **parallel** to the interface (perpendicular to both $d\boldsymbol{l}$ and $\hat{\boldsymbol{n}}$) so in a certain sense, the surface tension is a "one-dimensional pressure" that acts on any line along the surface of a liquid.

The total force exerted by the right region on the left is correspondingly

$$\boldsymbol{F} = \int_L \gamma d\boldsymbol{l} \times \hat{\boldsymbol{n}}, \tag{9.7}$$

where the integral is performed over the line L. The proof of this follows the previous procedure exactly. We can imagine cutting the liquid into two regions along L and exerting an external force $\boldsymbol{F}_{ext} = -\boldsymbol{F}$ on the left region to balance the force due to the right region. Applying the principle of virtual work to the left region when its entire boundary line with the right region stretches by an arbitrary displacement $\delta\boldsymbol{x}$ (while maintaining the same shape), the virtual work done by the external force is $\boldsymbol{F}_{ext} \cdot \delta\boldsymbol{x}$ while the virtual work done performed by the right region is negative of the change in surface energy of the right region. Since the change in surface area of the right region due to the displacement of an infinitesimal line segment $d\boldsymbol{l}$ by $\delta\boldsymbol{x}$ is $-\delta\boldsymbol{x} \cdot (d\boldsymbol{l} \times \hat{\boldsymbol{n}})$, the total increase in surface energy of the right region is

$$-\int_L \gamma \delta\boldsymbol{x} \cdot (d\boldsymbol{l} \times \hat{\boldsymbol{n}}).$$

Applying the principle of virtual work,

$$\boldsymbol{F}_{ext} \cdot \delta\boldsymbol{x} = -\int_L \gamma \delta\boldsymbol{x} \cdot (d\boldsymbol{l} \times \hat{\boldsymbol{n}}) = \delta\boldsymbol{x} \cdot -\int_L \gamma (d\boldsymbol{l} \times \hat{\boldsymbol{n}}).$$

Since this is valid for all $\delta\boldsymbol{x}$, we must have

$$\begin{aligned} \boldsymbol{F}_{ext} &= -\int_L \gamma d\boldsymbol{l} \times \hat{\boldsymbol{n}} \\ \implies \boldsymbol{F} &= \int_L \gamma d\boldsymbol{l} \times \hat{\boldsymbol{n}}. \end{aligned}$$

Excess Pressure in a Spherical Droplet

Surface tension causes a pressure discontinuity at a curved interface of a liquid at equilibrium. To start off, let us consider a spherical liquid droplet that has equilibrated in weightlessness. We wish to find the difference in pressure between the interior and the exterior of the drop, $p_1 - p_0$, given the surface tension γ and the radius of the drop r.

Method 1: Force We cut the sphere into two hemispheres. Considering the right hemisphere, we see that the force due to the difference between the

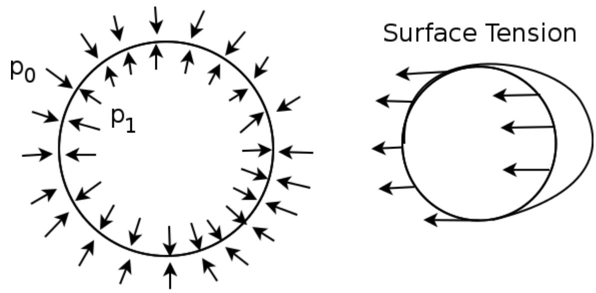


Figure 9.9: Liquid droplet

internal pressure and external pressure $p_1 - p_0$ tends to push it to the right. On the other hand, the surface tension (represented by arrows in the right of Fig. 9.9) balances that by pulling it to the left, towards the other hemisphere. In fact, evaluating the cross-product of the integrand in Eq. (9.7) along the perimeter of the equatorial circle would show that the surface tension force on the right hemisphere due to each line segment along the equatorial circle is exactly leftwards. Since the perimeter of the equatorial circle is $2\pi r$, the net surface tension force on the right hemisphere is $\gamma 2\pi r$. On the other hand, the force due to the pressure difference is simply[6] the pressure difference multiplied by the area of the equatorial circle πr^2. Thus, for the hemisphere to remain stationary,

$$(p_1 - p_0)\pi r^2 = 2\pi r\gamma$$

$$\Delta p = p_1 - p_0 = \frac{2\gamma}{r}.$$

Method 2: Virtual Work We shall apply the principle of virtual work to the molecules near the spherical surface. Suppose the sphere expands by a radius δr. Then, the virtual work done by the pressure on the surface molecules would be

$$\delta W_P = \Delta p 4\pi r^2 \delta r.$$

[6]To be completely rigorous, the force on an infinitesimal surface element $d\boldsymbol{A}$ (directed outwards) on the right hemisphere is $pd\boldsymbol{A}$ where p is the net pressure. Since there should only be a net force rightwards, we can simply take the rightwards component (defined to be aligned with the positive x-direction) of this, $pd\boldsymbol{A} \cdot \hat{\boldsymbol{i}}$, and integrate over the entire hemisphere to compute the net force. This becomes $\iint_{hemi} pd\boldsymbol{A} \cdot \hat{\boldsymbol{i}} = p \iint_{hemi} d\boldsymbol{A} \cdot \hat{\boldsymbol{i}}$. For an infinitesimal area $d\boldsymbol{A}$ that subtends an angle θ with $\hat{\boldsymbol{i}}$, the integrand evaluates to $dA \cos\theta$ which is simply the area of its projection onto the equatorial plane! Therefore, $\iint_{hemi} d\boldsymbol{A} \cdot \hat{\boldsymbol{i}}$ is simply the area of the projection of the hemisphere onto the equatorial plane which is just the area of the equatorial circle, πr^2. The net force is accordingly $p \cdot \pi r^2$.

The virtual work performed by surface tension can be computed as the negative change in surface energy. The increase in the area of the liquid surface can be calculated as follows. The surface area of a sphere is

$$A = 4\pi r^2$$
$$\implies \delta A = 8\pi r \delta r.$$

Thus, from the definition of surface tension,

$$\delta W_{ST} = -\delta U_{ST} = -\gamma \delta A = -8\pi r \gamma \delta r.$$

By the principle of virtual work, if the surface molecules were at equilibrium, the sum of all forms of virtual work must yield zero.

$$\delta W_P + \delta W_{ST} = 0 \implies \Delta p 4\pi r^2 \delta r = 8\pi r \gamma \delta r$$
$$\Delta p = p_1 - p_0 = \frac{2\gamma}{r}.$$

We see that the internal pressure of a liquid drop is in fact larger than the external pressure. As a word of caution, the last technicality in considering surface tension would be the number of interfaces. For example, if the liquid droplet in the above problem were to be replaced by a soap bubble, then

$$\Delta p = p_1 - p_0 = \frac{4\gamma}{r},$$

as there are now in fact, two liquid-vapor interfaces — the soap bubble is a hollow sphere with a liquid surface. Thus, there is a vapor-liquid-vapor transition from the outside to the inside of the soap bubble.

Problem: Referring to Fig. 9.10, two soap bubbles coalesce to become two spheres of radii r_1 and r_2 that are connected by a common soap interface that takes the form of the surface of a spherical cap with radius r_3 at equilibrium. Determine r_3.

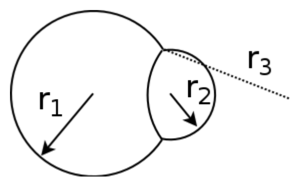


Figure 9.10: Coalesced bubbles (the case $r_2 < r_1$ is depicted above)

Let the external pressure be p_0. The pressures within the two soap bubbles at equilibrium are then

$$p_1 = p_0 + \frac{4\gamma}{r_1},$$
$$p_2 = p_0 + \frac{4\gamma}{r_2}.$$

Furthermore, the difference in the interior pressures of the soap bubbles must also be related to the radius of the spherical interface, i.e.

$$|p_1 - p_2| = \frac{4\gamma}{r_1 r_2}|r_2 - r_1| = \frac{4\gamma}{r_3}$$
$$\implies r_3 = \frac{r_1 r_2}{|r_2 - r_1|}.$$

The interface is concave with respect to the smaller sphere which has a larger interior pressure.

Young-Laplace Equation

For a general curved surface with a single interface at equilibrium, the Young-Laplace equation states that the difference in the internal and external pressures across the interface at a particular point P is

$$\Delta p = \gamma\left(\frac{1}{r_1} + \frac{1}{r_2}\right), \tag{9.8}$$

where r_1 and r_2 are the radii of curvature, which will be defined in the proof below, along two curves perpendicular at P, which are both located along the interface. The concave side of the interface has a larger pressure. One can easily check that by substituting $r_1 = r_2 = r$ for a sphere, the above equation reduces to what we have derived earlier.

Proof: Define two perpendicular directions along the relevant curved interface at a certain point P on the interface (these directions are perpendicular to the normal at P). Next, consider an infinitesimal rectangular surface element, centered about P, with side lengths dl_1 and dl_2 along those perpendicular directions.

Now, switch to the side view of this surface element where the side dl_2 is on this page and the side dl_1 is into the page (Fig. 9.11). In the first approximation, a curve in the vicinity of an arbitrary point P is locally approximated by a tangent line at P. However, this approximation is not

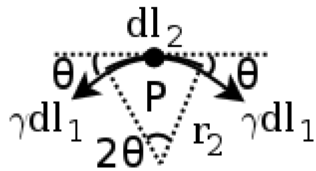


Figure 9.11: Side view

accurate enough in this context, as approximating dl_2 as a straight line would cause the surface tension forces on the sides of length dl_1 to nullify each other. We need a better approximation that better reflects the curved nature of this interface. To this end, we can approximate dl_2 locally by an arc of an appropriate circle that passes through P. Such a circle is known as an osculating circle[7] and its radius r is known as the radius of curvature of the curve at P. If the radius of curvature of dl_2 is r_2, the angle subtended by the surface tension forces along the sides with length dl_1 and the horizontal is $\theta \approx \frac{dl_2}{2r_2}$. The net force due to these forces is $2 \cdot \gamma dl_1 \cdot \frac{dl_2}{2r_2} = \frac{\gamma dl_1 dl_2}{r_2}$ downwards, thereby contributing to a downwards pressure $\frac{\gamma}{r_2}$ on the surface element.

Applying the same argument to the forces along the sides with length dl_2, the total pressure due to surface tension on this infinitesimal surface element is $\gamma(\frac{1}{r_1} + \frac{1}{r_2})$ downwards[8] — indicating that the pressure on the bottom must be larger than that on top by this amount to balance the pressure due to surface tension. We have hence proven our claim.

As an aside, observe that there are infinitely many pairs of perpendicular directions along the interface at P — such that r_1 and r_2 can be the radii of curvature along an arbitrary pair of curves along the interface that are perpendicular at P. However, we often choose r_1 and r_2 as a specific pair known as the principal radii of curvature which are defined as the maximum and minimum radii of curvature along all possible curves passing through P. It can be proven mathematically that these are the only extrema and that the directions of these curves, known as the principal directions, are mutually perpendicular at P. One can also derive the principal radii of curvature from the equation describing the surface, but we shall not delve further into

[7]Refer to any book on differential geometry to understand how the osculating circle of a point on a curve can be found.

[8]Note that even though we have drawn the surface to be concave with respect to the bottom, one can substitute a negative radius of curvature if the surface is convex with respect to the bottom.

the mathematical details as the radii of curvature of the surfaces that we encounter are usually obvious.

Problem: Determine the pressure discontinuity across a cylindrical interface of radius r and surface tension γ at equilibrium.

Consider a point P on the interface. Notice that the azimuthal and axial (parallel to the axis of the cylinder) directions are perpendicular at P with radii of curvature r and infinity — the latter because the center of a circle that approximates a straight line must be located at infinity.[9] Applying the Young-Laplace equation,

$$\Delta p = \frac{\gamma}{r}.$$

Solid-Liquid-Air Interface

Similar to the case of a liquid, we can define a surface energy for a solid surface in a vacuum and associate a surface tension γ_{sv} with the solid-vacuum interface.

When a liquid droplet is placed on a solid plate, there are really three interfaces — namely, the liquid-vapor, solid-vapor and solid-liquid interfaces. The former two can be described by the liquid-vacuum and solid-vacuum surface tensions γ_{lv} and γ_{sv} as we assume that the vapor does not interact with the other phases.

Things become more murky when we study the solid-liquid interface. We still ascribe a surface tension γ_{sl} to it but its meaning is subtly different. It represents the surface energy per unit area stored in the solid-liquid interface but it is not the external work required to increase the solid-liquid interface by a unit area by bringing molecules from the interior of the liquid to the solid-liquid interface. This becomes clear when we understand the origin of γ_{sl}. To create a solid-liquid interface, we can join a solid-vacuum interface and a liquid-vacuum interface which have surface energy densities γ_{sv} and γ_{lv} respectively. However, in this process, the attractive interactions between the solid and liquid molecules — which are known as adhesive interactions — are no longer negligible (unlike solid-vapor and liquid-vapor interactions). These adhesive interactions contribute to a negative potential energy as the solid-vacuum and liquid-vacuum interfaces are brought closer together — implying that the surface energy density γ_{sl} stored in the

[9] This is to ensure that an infinitesimal arc around P is straight.

solid-liquid interface formed is

$$\gamma_{sl} = \gamma_{sv} + \gamma_{lv} - A_{sl},$$

where A_{sl} is the external work per unit area required to separate the solid-liquid interface and revert to the solid-vacuum and liquid-vacuum interfaces. On the other hand, the external work needed to increase the solid-liquid interface by a unit area by bringing molecules within the liquid to the interface is[10] $\gamma_{lv} - A_{sl} = \gamma_{sl} - \gamma_{sv}$. In other words, the discrepancy between the surface energy density and the external work done per unit area arises from the adhesive interactions between the liquid and solid molecules, in addition to the cohesive interactions between molecules of the same kind.

The interplay of these cohesive and adhesive interactions determines how a liquid droplet responds when it is placed on a solid plate. In describing the shape of the droplet, the contact angle is defined as the angle subtended between the solid surface and the gradient of the fluid surface at the contact point on the solid-liquid interface. It is conventionally measured through the liquid and denotes the direction of the surface tension force (we mean along the liquid-vapour interface when we refer to just surface tension) on the liquid molecules near the contact point, due to the rest of the liquid (Fig. 9.12).

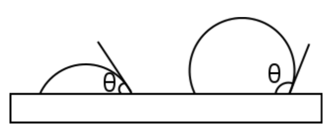


Figure 9.12: Contact angles on hydrophilic (left) and hydrophobic (right) surfaces

We can actually theoretically predict the equilibrium contact angle for a droplet placed on a plate. Since the droplet should be rotationally symmetric due to the infinite nature of the plane, a single cross-section is representative of the entire droplet. Consequently, consider the region in the vicinity of the contact point in a cross-section of a liquid droplet at equilibrium.

In Fig. 9.13, θ is the equilibrium contact angle. Suppose that we displace the contact point by a virtual distance δx outwards (towards the

[10] γ_{lv} is the external work (per unit area) required to extract the interior liquid molecules towards the interface if the exterior medium is vacuum. If the neighboring medium is instead a solid which engenders an attractive force on the molecules as they are brought in, we must subtract γ_{lv} by A_{sl} when computing the potential energy associated with the molecules, to account for the negative work done by the external force in overcoming these attractive interactions that tend to pull these molecules towards the solid.

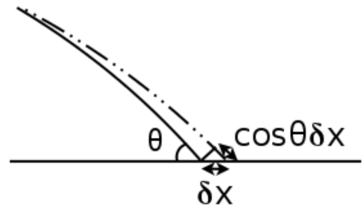


Figure 9.13: Displacement of contact point

vapor) while the shape of the liquid-vapor interface is roughly maintained due to the minuscule value of δx. The solid-vapor interface decreases by δx in length while the solid-liquid interface increases by δx in length. Meanwhile, the liquid-vapor interface increases by $\delta x \cos\theta$ in length. Therefore, in order for the sum of changes in potential energies to be zero (as the sum of all virtual works, due only to conservative forces in this case, must be zero),

$$\gamma_{sl}\delta x + \gamma_{lv}\cos\theta\delta x - \gamma_{sv}\delta x = 0$$
$$\implies \cos\theta = \frac{\gamma_{sv} - \gamma_{sl}}{\gamma_{lv}},$$

which is known as Young's equation. Now, we shall present an alternative derivation of this relationship from the perspective of forces. Firstly, we shall dismiss a common erroneous proof that is usually stated as follows.

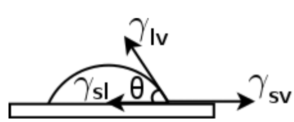


Figure 9.14: Surface tensions at contact point (incorrect argument)

Balancing the three forms of surface tension at the contact point along the horizontal direction in Fig. 9.14, we obtain

$$\gamma_{sv} = \gamma_{sl} + \gamma_{lv}\cos\theta$$
$$\implies \cos\theta = \frac{\gamma_{sv} - \gamma_{sl}}{\gamma_{lv}}.$$

Now, there are many fallacies in this argument. Firstly, we cannot balance the forces at the contact point as it is just a mathematical point that does not contain any real particles. Secondly, even if we consider the region of particles close to the contact point, what type of particles are we encompassing in our system? For instance, γ_{sv} refers to the cohesive forces on solid molecules while γ_{lv} refers to the cohesive forces on liquid molecules! Thirdly, why is the vertical component of forces unbalanced?

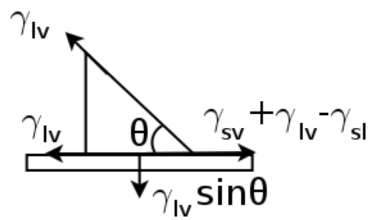


Figure 9.15: Forces on wedge (correct argument)

To rectify these loopholes, it is best to explicitly define the system that we consider from the start. We choose the liquid molecules near the contact point, which take the form of the wedge above, as our system. The forces on this system are the cohesive forces due to the rest of the liquid, described by the liquid-vacuum/vapor surface tension γ_{lv} on the hypotenuse and the base, and the force due to the solid (the interactions between the liquid and vapour are negligible). The force on this system in the plane of the plate, due to the solid, is easy to compute as we have already proven that the work per unit area due to the solid-liquid adhesion is $A_{sl} = \gamma_{sv} + \gamma_{lv} - \gamma_{sl}$. Therefore, the adhesive force per unit length on our system in the horizontal direction must be A_{sl} outwards (towards increasing interfacial area). Note that all surface tension forces are due to the infinitesimal segment along the interface pointing into the page.

Referring to Fig. 9.14, the cohesive surface tension force per unit length γ_{lv} along the liquid-vapor interface on the molecules in the vicinity of the contact point is drawn correctly. However, the γ_{sv} and γ_{sl} forces per unit length should instead be reflected as a $A_{sl} = \gamma_{sv} + \gamma_{lv} - \gamma_{sl}$ force per unit length rightwards, produced by the solid, and a γ_{lv} force per unit length leftwards by the rest of the fluid due to the surface tension on the solid-liquid interface.

Finally, to correct the seemingly-unbalanced vertical component of γ_{sv}, we note that we have only computed the force on our system due to the solid in the plane of the plate. There are still[11] forces of adhesion between the liquid and solid molecules near the interface which balance the upwards component of γ_{lv}. The correct diagram is thus Fig. 9.15. Having clarified

[11]By considering infinitesimal elements in the interior of the fluid, one can easily show that this normal component of adhesive forces $\gamma_{lv}\sin\theta$ is strictly localized to the three-phase line.

these facts, we can now safely write

$$\gamma_{sv} - \gamma_{sl} = \gamma_{lv} \cos\theta$$
$$\implies \cos\theta = \frac{\gamma_{sv} - \gamma_{sl}}{\gamma_{lv}}.$$

Furthermore, we can see that the force per unit length exerted by the solid on the liquid, in the plane of the solid-liquid interface, is in fact

$$A_{sl} = \gamma_{sv} + \gamma_{lv} - \gamma_{sl} = \gamma_{lv}(1 + \cos\theta).$$

There is a moral to be told here. It is often advised to adopt a thermodynamic approach (finding changes in energy and applying the principle of virtual work) instead of a mechanical one when dealing with surface tension. The latter is much more subtle and is thus error-prone, especially when we are confused about the system that we are referring to.

Now, there are a few regimes of interest that determine the response of the liquid droplet (refer back to Fig. 9.12). Firstly, if $0 < \gamma_{sv} - \gamma_{sl} \le \gamma_{lv}$, then $0 \le \theta < \frac{\pi}{2}$. The surface is said to be mostly wetting (hydrophilic) and the liquid droplet equilibrates with an acute contact angle. On the other hand, if $0 < \gamma_{sl} - \gamma_{sv} < \gamma_{lv}$, we find that $\frac{\pi}{2} < \theta < \pi$ and the surface is described as mostly non-wetting (hydrophobic). The droplet equilibrates with an obtuse contact angle. If $0 < \gamma_{lv} < \gamma_{sv} - \gamma_{sl}$, there are no solutions for θ. Instead, an equilibrium does not exist as the outwards force due to the solid is so strong that the liquid is dragged further into the vapor region — spreading the liquid over the entire solid surface. In an apt manner, the surface is said to be completely wetting. Similarly, if $0 < \gamma_{lv} < \gamma_{sl} - \gamma_{sv}$, the inwards force due to the solid is so strong that the droplet is dragged into the liquid-region and is thus contracted. In fact, it is more energetically favorable to introduce vapor films between portions of the liquid such that the liquid is actually collected into copious spherical pearls. Such a surface is said to be completely non-wetting.

Capillary Action

The most vivid manifestation of the various surface tensions is the phenomenon of capillary action. When a tube with a small cross-section is placed in a container of liquid, the liquid is "sucked" into the tube and rises above the liquid level in the container.

Suppose that the tube is a cylinder of small radius r and that the equilibrium height of the liquid in the vertical column is h above the liquid level in the container (we neglect the volume of the meniscus). The liquid has density ρ. We shall present three arguments to determine h in terms of the liquid-vacuum surface tension, that we will just write as γ, and the contact angle θ.

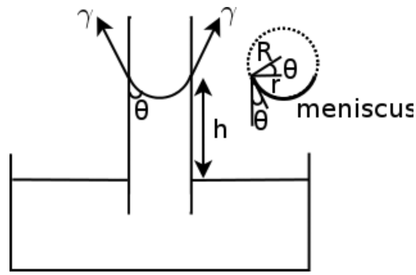


Figure 9.16: Capillary action

Firstly, suppose that the height of the liquid within the tube increases by δh. Since the height of the solid-liquid interface is increased by δh while that of the solid-vapor interface is decreased by δh, the change in surface energy is $(\gamma_{sl} - \gamma_{sv})2\pi r\delta h$ where $2\pi r$ is the perimeter of the cross-sectional circle of the tube. There is no change in surface energy associated with the liquid-vapor interface since its shape does not change. Meanwhile, there is also a change in gravitational potential energy as δh height of liquid (which corresponds to $\rho\pi r^2\delta h$ of mass) is effectively transferred from the bottom of the tube to an altitude h — indicating an increase in gravitational potential energy by $\rho\pi r^2 gh\delta h$. Equating the sum of the changes in potential energies to zero, by the principle of virtual work,

$$(\gamma_{sl} - \gamma_{sv})2\pi r\delta h + \rho\pi r^2 gh\delta h = 0$$
$$h = \frac{2(\gamma_{sv} - \gamma_{sl})}{\rho rg} = \frac{2\gamma\cos\theta}{\rho rg},$$

where we have applied Young's equation in the last step. Note that Young's equation is still valid even if gravity is now parallel to the surface of the plate, as the weight of the liquid molecules in the vicinity of a contact point is negligible compared to the other forces they experience.

Moving on to the second method, we can impose the condition that the liquid pressure must be uniform throughout a vertical level at hydrostatic equilibrium. Let p_0 be the atmospheric pressure. As r is small, the meniscus can be treated as a spherical cap with radius $R = \frac{r}{\cos\theta}$ (see Fig. 9.16). Therefore, the pressure in the liquid directly below the meniscus is $p_0 - \frac{2\gamma\cos\theta}{r}$ by the Young-Laplace relationship. Since the pressure at the water level inside the open container must be p_0, we have

$$p_0 - \frac{2\gamma\cos\theta}{r} + \rho gh = p_0$$
$$h = \frac{2\gamma\cos\theta}{\rho rg}.$$

Finally, we can also solve this problem by analyzing the forces on the portion of water in the tube above the water level in the container. Firstly, we know from the previous section that the force per unit length that the tube exerts on the liquid at the three-phase interface, along the surface of the tube (towards the vapor), is $A_{sl} = \gamma(1+\cos\theta)$. Therefore, the upwards force exerted by the tube on the liquid is $2\pi r\gamma(1+\cos\theta)$. Next, there is also a downwards force due to surface tension on the bottom boundary circle of this system, $2\pi r\gamma$, that is exerted by the water beneath it. Observing that there is no pressure difference between the top and bottom ends of this water column, the net effect of these forces must balance the weight of the water in our system.

$$2\gamma\pi r\cos\theta = \rho\pi r^2 gh$$
$$h = \frac{2\gamma\cos\theta}{\rho r g}.$$

Evidently, the height of the liquid column increases with a smaller tube, ceteris paribus. Finally, also note that this result is valid for both acute and obtuse contact angles even though we drew the acute case in the diagram. The liquid level in the tube will be lower than the liquid level in the container if the contact angle is obtuse.

Problem: Determine the depth H of a meniscus (i.e. the height of the highest level minus the height of the lowest level) in a cylindrical tube with an arbitrary radius that contains a liquid of density ρ, surface tension γ and contact angle θ. You can no longer assume that the shape of the meniscus is a spherical cap in this case, but you may assume that the slope of the meniscus is small everywhere. With your result, justify why we could ignore the volume of the meniscus for small tube radii previously.

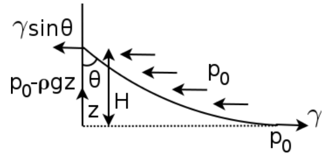


Figure 9.17: Half of meniscus in cross-section

Again, since the tube is axially-symmetric, we can consider a single cross-section with half of the meniscus. Assuming that the contact angle is acute, the highest liquid level should occur at the interface with the tube while the lowest liquid level should occur at the center of the meniscus.

In fact, the center of the meniscus must be completely flat — else there is no chance for the surface tension on the molecules near the center to be balanced in the vertical direction. This flatness implies that the pressure in the liquid, directly below the meniscus, is the atmospheric pressure p_0 by the Young-Laplace equation (with infinite radii of curvature). Therefore, the hydrostatic pressure at a point on the wall of the tube must be $p_0 - \rho g z$, where z is the point's height above the center of the meniscus.

Now, we can determine the depth H of the meniscus by balancing the horizontal forces on the half of the meniscus depicted on the previous page. Firstly, the liquid experiences a uniform atmospheric pressure p_0 from its right — resulting in a net leftwards force per unit length $p_0 H$ in this cross-section (we can take the force to be purely horizontal because the slope is small). Next, since the liquid exerts a hydrostatic pressure on the wall, it must also experience a pressure due to the wall. As the hydrostatic pressure varies linearly with depth, we can simply take the average of the hydrostatic pressure multiplied by the depth of the liquid to compute the net force per unit length on the liquid. This amounts to $(p_0 - \frac{\rho g H}{2}) \cdot H = p_0 H - \frac{\rho g H^2}{2}$ force per unit length on the liquid rightwards. Besides the two pressures on the liquid, there are also the surface tension at the center of the meniscus which delivers a rightwards force per unit length γ, and the horizontal (normal) attractive force per unit length exerted by the tube on the liquid at the three-phase interface, $\gamma \sin\theta$, as discussed previously (to balance the normal component of surface tension due to the rest of the fluid). Balancing the forces in the horizontal direction while taking rightwards to be positive,

$$p_0 H - \frac{\rho g H^2}{2} - p_0 H + \gamma - \gamma \sin\theta = 0$$

$$H = \sqrt{\frac{2\gamma(1 - \sin\theta)}{\rho g}},$$

which is surprisingly, independent of the tube radius. One can use a similar procedure to show that the above is also valid for obtuse contact angles (the center of the meniscus is now higher than the liquid level near the tube). Comparing H with the expression for h derived previously, we observe that $H \ll h$ if $r \ll \sqrt{\frac{\gamma}{\rho g}}$ (we are now able to quantify what we mean by a small tube radius)! In fact, the expression $\sqrt{\frac{\gamma}{\rho g}}$ is defined as the capillary length L_c, which is a characteristic length scale for the interface between two media, so we can say that our answer for h is valid when $r \ll L_c$.

Floating Objects

In the previous section, we studied the forces on a liquid due to a solid — we will focus on the reverse in this section. There are three main effects that the presence of a liquid induces on the solid. Firstly, there is an upthrust on a submerged or partially submerged body that is generated due to the varying pressure in water with height. Secondly, the liquid molecules along the three-phase interface also exert the reaction pair to the adhesive forces due to the solid. This results in forces per unit length $\gamma(1+\cos\theta)$ tangential to the solid surface (pointing towards the liquid) and $\gamma\sin\theta$ normal to the solid surface (also pointing towards the liquid) on the solid, where θ is the contact angle. Finally, there is generally an additional "curvature pressure", superimposed on the normal pressure variation with height, due to the local curvature of the solid-liquid interface. The origin of this is simple — as we have accounted for the adhesive interactions, the remaining effect of the solid is akin to vacuum. If the solid-liquid interface is concave with respect to the solid, the solid has to exert an additional Laplace pressure to keep the liquid molecules near the interface in equilibrium (via a normal force). The liquid molecules then exert the reaction pair to this normal force which effectively constitutes an additional curvature pressure.[12] Therefore, the curvature pressure on the solid is positive (acting towards the solid) if the solid-liquid interface is convex with respect to the liquid. The converse statement holds for the opposite shape of the interface.

The combination of the two latter factors can actually help non-wetting solid objects, such as water beetles, to float on a less-dense liquid! Let us consider the simplest example of a sphere floating on water.

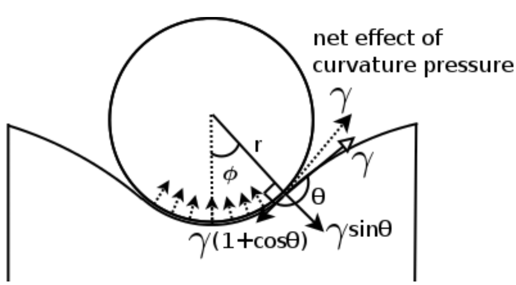


Figure 9.18: Floating sphere

In the figure above, a sphere of radius r is floating with an angle of contact θ. ϕ is the angular coordinate of the three-phase (solid, liquid, vapor)

[12]Another perspective is that the Laplace pressure enhances the pressure in the liquid, beyond the usual linear variation with depth which is associated with the upthrust.

intersection in a cross-section and corresponds to the location where the reaction to the adhesive forces act ($\gamma(1 + \cos\theta)$ tangentially and $\gamma \sin\theta$ normally with respect to the sphere as shown above). Now, the curvature of the sphere leads to a uniform curvature pressure $\frac{2\gamma}{r}$ along the entire solid-liquid interface (depicted by the dashed arrows). This amounts to an upwards force of magnitude $\frac{2\gamma}{r} \cdot \pi r^2 \sin^2\phi = 2\gamma\pi r \sin^2\phi$ (pressure multiplied by the area of the three-phase boundary circle which has radius $r \sin\phi$). Observe that this is equivalent to saying that the net effect of the curvature pressure is to introduce a tangential (relative to the sphere) surface tension force γ on the circle at the three-phase interface (pointing upwards) as it will act on a perimeter $2\pi r \sin\phi$ for a total upwards force of $\gamma \cdot 2\pi r \sin\phi \cdot \sin\phi = 2\gamma\pi r \sin^2\phi$ (multiply by $\sin\phi$ to extract the vertical component).

Now, let's combine the effects of the curvature pressure and the adhesive forces to obtain forces per unit length $\gamma \cos\theta$ tangential to the sphere and $\gamma \sin\theta$ normal to the sphere, both pointing towards the liquid, along the three-phase interface. This implies that the net effect is simply a force per unit length γ along the three-phase intersection, tangential to the liquid surface (white arrow)! With this remarkable result, we can compute the force due to this net effect as

$$F = \gamma \cdot 2\pi r \sin\theta \cdot \sin\left(\frac{3\pi}{4} - \theta - \phi\right) = 2\gamma\pi r \sin\theta \sin\left(\frac{3\pi}{4} - \theta - \phi\right),$$

where we multiply by $\sin(\frac{3\pi}{4} - \theta - \phi)$ to retrieve the upwards component. The total force balance for the sphere is thus

$$mg = F_{upthrust} + 2\gamma\pi r \sin\theta \sin\left(\frac{3\pi}{4} - \theta - \phi\right),$$

where m is the mass of the sphere and $F_{upthrust}$ is the upthrust that it experiences. The combination of the effects associated with surface tension and upthrust helps the sphere to stay afloat. Now, you may wonder if the presence of adhesive interactions and curvature pressure invalidates the previous proof of Archimedes' Principle which relied on the fluid's inability to distinguish the entity that it exerts a force on. To ease your worries, observe that no such forces exist on a line of molecules when the entity is in fact an identical portion of fluid — implying that the previous proof is still valid. Thus, the expression for the upthrust still remains the same but one has to account for the additional effects due to the other factors.

Finally, the above result is in fact completely general. For an object of an arbitrary shape, the net vertical force due to adhesive interactions and curvature pressure is akin to introducing a force per unit length γ at the three-phase interface, along the surface of the fluid! Therefore, if we place a thin cylinder of mass m, length l and negligible radius in a liquid with surface tension γ and contact angle of π radians, the force balance equation is $2\gamma l = mg$ (the upthrust is negligible). However, be wary that no such short-cut exists for the components of force along the plane perpendicular to the vertical — one has to account for the adhesive forces and curvature pressure manually.

9.4 Fluid Dynamics

The flow of a fluid can be divided into several categories. A uniform flow is one where the flow characteristics — such as pressure and flow velocity — are identical at all points. A steady flow is one whose flow characteristics do not vary over time but is not necessarily uniform. Lastly, an incompressible flow occurs when the density of the fluid is identical everywhere.

This section will analyze steady and incompressible one-dimensional flows of fluids. In our set-ups, cross-sections are assumed to be small such that the flow characteristics — such as velocity — are uniform about a cross-section. We shall start with a few definitions. A streamline is a line that is tangential to the instantaneous flow velocity at all points along the streamline (akin to an electric field line) and reflects the trajectory of a fluid particle lying on that streamline if the flow is steady. Now, consider the set of all streamlines that pass through a closed loop (such as the left loop in the figure below). Such a collection of streamlines is known as a streamtube (e.g. the tube that looks like a pipe in the figure below). Since streamlines cannot cross (as that would imply two directions of flow velocity at the points of intersection), fluid particles can only flow along the cross-section of a streamtube and cannot escape by cutting across the surface of a streamtube.

The Mass Continuity Equation

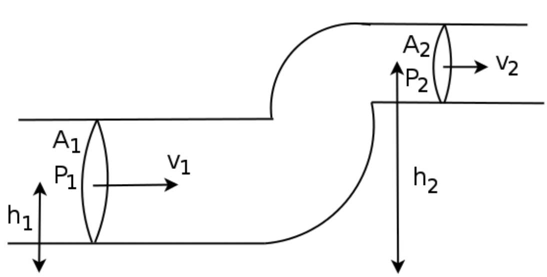


Figure 9.19: Narrow streamtube (the above is not a physical pipe)

In a steady-state system, the total mass within a section of fluid must be invariant over time. Therefore, if we define ρ_1, A_1, v_1 to be the mass density, cross-sectional area and flow velocity at one section of a narrow streamtube (narrow such that these properties are uniform over a cross-section) and ρ_2, A_2, v_2 to be those at another section (Fig. 9.19), we have

$$\rho_1 A_1 v_1 = \rho_2 A_2 v_2. \tag{9.9}$$

The expression $\rho A v$ is the mass flux, which is the rate of mass flow through a cross-section. This must be equal at all sections of the streamtube as whatever goes in must come out at the same rate. If this were not the case, mass would be accumulated along the flow and the flow would no longer be steady. Furthermore, in the case of incompressible flow, $\rho_1 = \rho_2$. Then,

$$Av = Q \tag{9.10}$$

for a constant Q which is known as the volume flow rate.

Bernoulli's Equation

For an incompressible, steady and energy-conserving one-dimensional fluid flow, Bernoulli's principle states that for two points along a streamline or a narrow streamtube,

$$P_1 + \frac{1}{2}\rho v_1^2 + \rho g h_1 = P_2 + \frac{1}{2}\rho v_2^2 + \rho g h_2, \tag{9.11}$$

where P is the pressure at a point along the fluid flow, v is the flow velocity at that point and h is its height with respect to a certain reference point. The dimensions of the cross-sections of the streamtube of concern are assumed to be negligible compared to their heights, such that the gravitational potential energies of all mass elements on a particular cross-section are assumed to be identical.

Proof: Consider the evolution of the portion of fluid enclosed between the two cross-sections in Fig. 9.19 after a time interval dt. $\rho Q dt$ mass of fluid would have advanced a distance $v_1 dt$ from the left cross-section and $v_2 dt$ from the right cross-section. Thus, the work done by the pressure at the two ends is

$$W = (P_1 - P_2) Q dt.$$

Furthermore, the velocity of $\rho Q dt$ mass of fluid effectively changes from v_1 to v_2. Therefore, the change in kinetic energy is

$$\Delta T = \frac{1}{2}\rho Q v_2^2 dt - \frac{1}{2}\rho Q v_1^2 dt.$$

Finally, the change in gravitational potential energy is evidently

$$\Delta U = \rho g Q dt(h_2 - h_1).$$

By the work-energy theorem,

$$W_{noncon} = \Delta(T + U),$$

where $W_{noncon} = (P_1 - P_2)Qdt$ in this case. Canceling the Qdt terms,

$$P_1 - P_2 = \frac{1}{2}\rho v_2^2 - \frac{1}{2}\rho v_1^2 + \rho g h_2 - \rho g h_1$$

$$P_1 + \frac{1}{2}\rho v_1^2 + \rho g h_1 = P_2 + \frac{1}{2}\rho v_2^2 + \rho g h_2.$$

Problem: An open cylindrical container of height h_0 and cross-sectional area A_1 is initially filled with water. There is a tap at the bottom of the container with cross-sectional area A_2 and negligible height. Assuming energy-conserving flow, find the height of the water level $h(t)$ as a function of time.

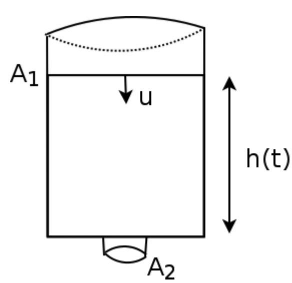


Figure 9.20: Water tap

Let u be the velocity of the water level and let v be the velocity of the water flushing through the tap, both defined to be positive downwards. By Bernoulli's equation,

$$\frac{1}{2}u^2 + gh = \frac{1}{2}v^2,$$

where we have canceled the atmospheric pressure terms on both sides. Furthermore, the rate of volume flow outwards is $A_2 v$. Thus,

$$\frac{dV}{dt} = -A_2 v.$$

Since $V = A_1 h$ and $u = -\frac{dh}{dt}$,

$$A_1 u = A_2 v \implies v = \frac{A_1}{A_2}u$$

which is effectively the mass continuity equation. Substituting this expression for v into the first equation,

$$\frac{1}{2}u^2 + gh = \frac{1}{2}\frac{A_1^2 u^2}{A_2^2}$$

$$\left(\frac{A_1^2}{A_2^2} - 1\right)u^2 = 2gh.$$

Substituting $u = -\frac{dh}{dt}$,

$$\frac{dh}{dt} = -\sqrt{\frac{2gA_2^2 h}{A_1^2 - A_2^2}}.$$

Separating variables and integrating,

$$\int_{h_0}^{h} \frac{1}{\sqrt{h}}dh = \int_0^t -\sqrt{\frac{2gA_2^2}{A_1^2 - A_2^2}}dt$$

$$2\sqrt{h} - 2\sqrt{h_0} = -\sqrt{\frac{2gA_2^2}{A_1^2 - A_2^2}}t$$

$$h = \left(\sqrt{h_0} - \sqrt{\frac{gA_2^2}{2\left(A_1^2 - A_2^2\right)}}t\right)^2,$$

which is only valid for $t \leq \sqrt{\frac{2(A_1^2 - A_2^2)h_0}{gA_2^2}}$. Note that this result is definitely inaccurate in the regime where h is small as the water surface may begin to shrink in radius (in addition to a decreasing fluid level) — introducing new complications.

Systems with a Small Aperture

In the case of containers with small apertures, the fluid inside the container can be taken to be approximately static, such that its squared speed in the application of Bernoulli's principle can be neglected (as it is second-order). However, the velocity of the ejected fluid and the velocity of the fluid inside the container when applying the continuity equation (which is first-order) are non-negligible.

Problem: A hollow cylinder of base area A and length l is initially filled with a fluid of mass density ρ. It lies with its cylindrical axis along the plane of a horizontal table. Then, a massless piston is pushed with a constant force

F at one of the ends of the fluid. If a small aperture of area $A_0 \ll A$ is made at the other end of the cylinder, determine the time required to empty the cylinder. Neglect atmospheric pressure.

Applying Bernoulli's principle to the end of the fluid that is touching the piston and the point directly outside the aperture,

$$P = \frac{1}{2}\rho v^2,$$

where $P = \frac{F}{A}$ is the pressure of the fluid at the end with the piston and v is the velocity of the ejected fluid (the velocity of the fluid in the cylinder is negligible). Thus, the volume flow rate is

$$Q = A_0 v = A_0\sqrt{\frac{2F}{\rho A}}.$$

The time required to empty the container is then

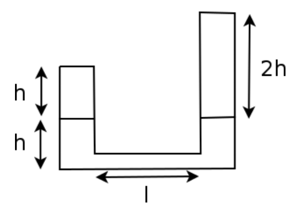
$$t = \frac{Al}{Q} = \frac{Al\sqrt{\rho A}}{A_0\sqrt{2F}}.$$

Problems

Pressure and Surface Tension

1. ***Accelerating Tube****

Consider the tube below with a constant horizontal cross-section — the length of the bends can be neglected. The ends are initially uncapped and the water level is even as the tube is stationary. The tube is capped and accelerated. If the lengths of the left and right air columns are h and $2h$ initially and $\frac{3}{2}h$ and $\frac{3}{2}h$ afterwards, determine the acceleration of the tube. Assume that the ideal gases undergo an isothermal process. Denote ρ as the density of water and p_0 as atmospheric pressure.



2. ***Falling Drop****

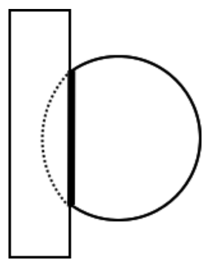
Assume that raindrops are spherical with a constant density ρ. Suppose that a raindrop of radius $\sqrt[3]{7}R$ falls from a height h and coalesces instantaneously with a raindrop of radius R. If the surface tension of a raindrop is γ and its specific heat capacity is c, determine the temperature rise immediately after the collision.

3. ***Colliding Soap Bubbles*****

The surface tension of soap bubbles in vacuum is γ. Now, two spherical soap bubbles are in equilibrium with respective radii r_a and r_b in an environment where the external pressure is zero. The two soap bubbles then coalesce to form a spherical soap bubble of equilibrium radius r_c, without any heat transfer between it and the environment. If the initial temperatures of the soap bubbles are both some unknown temperature T, show that the final equilibrium temperature of the combined soap bubble is also T. Determine r_c.

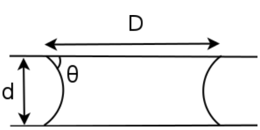
4. ***Balloon*****

A spherical balloon of total mass m, radius R and surface tension γ is tossed at a wall. The balloon then undergoes a small deformation of the form shown in the figure below (a portion is flattened). However, the part that has yet to collide with the wall is still rigid. If the balloon remains mostly spherical such that the pressure inside is uniform and unchanged, show that the motion of the balloon is simple-harmonic. Thus, determine the time between the initial collision and the rebound. (Estonian-Finnish Olympiad)



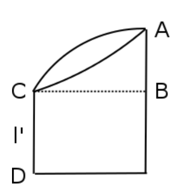
5. *Capillary Forces between Parallel Plates***

A thin film of water with surface tension γ and contact angle θ rests between two large parallel plates. The diameter of the circular liquid-solid interface on a disk is D while the distance between the plates is $d \ll D$. Neglecting gravity, determine the force experienced by each of the plates.



6. *Accelerating Cylinder***

A stationary cylinder of radius R contains a fluid of uniform density of height l. Now, the cylinder is given a constant leftwards acceleration a. If the resultant equilibrium shape of the fluid is the cylindrical segment shown below, determine the height l' of the cylinder, obtained by truncating the excess portion of the cylindrical segment. What is the maximum value of a for which the fluid covers the entire base of the cylinder?



7. *Ball in Hole***

A circular hole of radius r at the bottom of a tank of water with density ρ (the lid is open) is sealed by a sphere of radius $R > r$ and mass m. The water level in the tank is reduced to a certain height h, at which point the ball starts to rise out of the hole. Determine h while assuming that the water level in the tank is still above the top of the ball at the juncture where it just begins to rise.

Buoyancy

8. *Melting Ice**

An ice cube floats on water. Describe qualitatively the changes in the height of the fluid level when the ice melts in the 3 following cases: (1) A metal bead is embedded inside the ice cube. (2) A solid, whose density is less than that of water, is embedded. (3) The ice cube traps oil which is immiscible with water.

9. *Rock in Bowl**

A rock of mass m_r and density ρ_r is placed into a bowl of mass m_b. This system is then floated in a beaker of water with density ρ_w and cross-sectional area A — causing the water level to rise by height h_1. Subsequently, the rock is removed from the bowl and dropped into the beaker such that the final water level is still higher than the original water level by h_2. Determine $h_1 - h_2$.

10. *Moving Ants**

A group of ants are now trapped in an inverted, massless equilateral triangle of side length l and width w. The equilateral triangle is completely submerged in a fluid of density ρ. Determine the largest mass of ants that can rest at the top of the equilateral triangle such that the system is in stable equilibrium.

11. *Oscillation Between Two Fluids***

Suppose that a cylinder of mass m, cross-sectional area A and length l is in a state of equilibrium at the interface of two immiscible fluids. The cylindrical axis is along the vertical. If the densities of the top and bottom fluids are ρ_1 and ρ_2 respectively, show that the system can undergo simple-harmonic motion under certain conditions. Determine the angular frequency in such situations.

Fluid Dynamics

In the following questions, assume steady, incompressible and energy-conserving flow unless stated otherwise. Also assume that the flow characteristics are uniform about a cross-section by default.

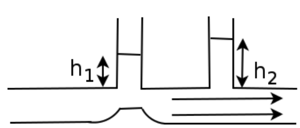
12. *Spraying Water**

An uncapped container with a large cross-sectional area currently holds water of height H above the ground. Determine the height h at which a small hole should be made such that the ejected water attains the greatest horizontal range on the ground.

Next, suppose that we poke many small holes on the container along the same vertical line, at various heights. The envelope of the water trajectories is defined as follows. For every vertical height level, we can find the point on a trajectory which is the furthest horizontal distance away from the container, at that height level. The envelope is then formed by the locus of such points. Determine the shape of the envelope while assuming that the water level in the container does not vary by much.

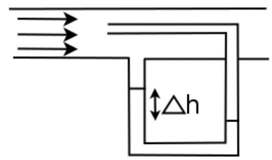
13. *Venturi-meter**

Consider the set-up below. Water flows from the left to right. If the heights of the water levels of the two tubes are h_1 and h_2 respectively and if the cross-sectional areas immediately below these pipes are A_1 and A_2, determine the volume flow rate in the pipe.



14. *Pitot Tube**

A vessel is currently carrying some air of constant density ρ_a. The air is traveling towards the right at a velocity v. Now, a pitot tube is inserted as shown in the figure on the next page such that the gas in the tube is stationary. If the fluid in the pitot tube has density $\rho_l \gg \rho_a$ and the difference in fluid levels is Δh, determine v.

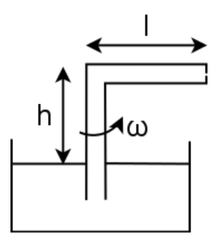


15. *Filling a Tank**

After a tap above an empty rectangular tank has been opened, the tank is filled with water at a constant rate in time T_1. After the tap has been closed, poking a small hole at the bottom of the tank empties it in another duration of time T_2. If the tap above the tank is now opened again, for what ratios of $\frac{T_1}{T_2}$ will the tank overflow?

16. *Sucking Water**

A tube of height h and length l is spinning in a beaker of water of density ρ at a constant angular velocity ω. If a small aperture of area A is punctured at the top-right end of the tube, determine the velocity of the emitted fluid at the aperture in the rotating frame of the tube and lab frame. Determine the external power required to maintain such a system.



17. *Transmitting Water**

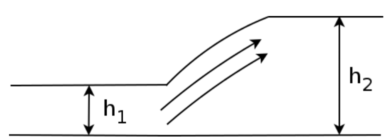
Two uncapped containers are placed on a level ground, right next to each other. The cross-sectional areas of the containers are A_1 and A_2 respectively which are both large. Initially, the containers carry water of constant density ρ. The water level of the first container is x_0 higher than that of the second. Now, a hole of area A ($A \ll A_1$ and $A \ll A_2$) is poked at the bottoms of both containers such that water flows from one container to another through the hole. Determine the time required for this system to equilibrate.

18. *Curving Water***

A massive pipe rests on a horizontal table. In the pipe, a fluid of constant density ρ flows from left to right with an angle of deflection θ, as shown in the figure below. Suppose that the areas of the left and right ends are A_1 and A_2 respectively with corresponding pressures of P_1 and P_2. Determine the force exerted on this portion of fluid by the pipe.

19. *Hydraulic Jump***

A fluid of density ρ undergoes steady, incompressible flow with an average flow velocity v_1 at the left end of the figure, towards the right end. It undergoes a hydraulic jump from an initial height h_1 to a final height h_2. If the system has a width w directed into the page and is uniform about this width, determine h_2 by the impulse-momentum theorem. Note that energy is not conserved and the flow characteristics are not uniform about a cross-section. Neglect atmospheric pressure.



Solutions

1. Accelerating Tube*

The initial pressures of the gases, immediately after the ends are capped, are p_0, the atmospheric pressure. As the volumes of the left and right columns increased by a factor of $\frac{3}{2}$ and decreased by a factor of $\frac{3}{4}$ respectively, the final pressure in the left and right columns are $\frac{2}{3}p_0$ and $\frac{4}{3}p_0$. Now, consider the horizontal portion of the tube which has a constant cross-sectional area A. The net force on this portion is the difference in pressure at its ends, multiplied by A.

$$F = \left(\frac{4}{3}p_0 + \rho gh - \frac{2}{3}p_0\right) A = \left(\frac{2}{3}p_0 + \rho gh\right) A$$

leftwards, where ρgh is the pressure difference due to the difference in liquid levels ($\frac{1}{2}h$ as compared to $\frac{3}{2}h$). This must be equal to the mass of the middle portion, multiplied by the acceleration a, so that

$$F = \rho Ala.$$

Then,

$$a = \frac{2p_0 + 3\rho gh}{3\rho l}$$

leftwards.

2. Falling Drop*

Let the masses of the raindrops be $7m$ and m respectively. The velocity of the first drop before the collision is $v = \sqrt{2gh}$. The velocity after the collision is given by the conservation of momentum to be

$$v' = \frac{7}{8}v.$$

The loss in kinetic energy is

$$\Delta T = \frac{1}{2}(7m + m)v'^2 - \frac{1}{2} \cdot 7mv^2 = -\frac{7}{16}mv^2.$$

There is also a change in surface energy due to the change in the total surface area of the bubbles when they coalesce. The final radius of the drop is, by the conservation of mass,

$$R' = \sqrt[3]{7R^3 + R^3} = 2R.$$

Therefore, the change in surface energy is

$$\Delta U = 4\pi\gamma(R'^2 - \sqrt[3]{49}R^2 - R^2) = 4\pi\gamma(3 - \sqrt[3]{49})R^2.$$

The heat evolved is

$$\Delta Q = -\Delta T - \Delta U.$$

The change is temperature is then

$$\Delta t = \frac{\Delta Q}{8mc} = \frac{7gh}{64c} + \frac{3\gamma(\sqrt[3]{49} - 3)}{8\rho c R}.$$

3. Colliding Soap Bubbles**

Referring to the section regarding the pressure discontinuity across a spherical interface, the initial pressures inside the soap bubbles are, respectively,

$$p_a = \frac{4\gamma}{r_a},$$
$$p_b = \frac{4\gamma}{r_b}$$

where we remember that there are two interfaces for a soap bubble. Therefore, by the ideal gas law, the number of moles of gas in the soap bubbles is

$$n_a = \frac{p_a V_a}{RT} = \frac{\frac{4\gamma}{r_a} \cdot \frac{4\pi r_a^3}{3}}{RT} = \frac{16\gamma\pi r_a^2}{3RT},$$
$$n_b = \frac{16\gamma\pi r_b^2}{3RT}.$$

Let the final equilibrium temperature of the combined soap bubble be T'. Then, for the total number of moles to be constant,

$$\frac{16\gamma\pi r_a^2}{3RT} + \frac{16\gamma\pi r_b^2}{3RT} = \frac{16\gamma\pi r_c^2}{3RT'}$$
$$\frac{r_a^2 + r_b^2}{T} = \frac{r_c^2}{T'}.$$

Furthermore, by the conservation of energy, the sum of the internal energy of the gases and the surface energy must be a constant.

$$\frac{3}{2}(n_a + n_b)RT + \gamma 4\pi(r_a^2 + r_b^2) = \frac{3}{2}(n_a + n_b)RT' + \gamma 4\pi r_c^2.$$

Substituting the expressions for n_a and n_b and rearranging,

$$\frac{\gamma 8\pi}{T}(r_a^2 + r_b^2)(T' - T) = \gamma 4\pi(r_a^2 + r_b^2 - r_c^2).$$

Substituting $r_c^2 = (r_a^2 + r_b^2)\frac{T'}{T}$ yields

$$(r_a^2 + r_b^2)(T' - T) = 0$$
$$\implies T' = T.$$

Substituting $T' = T$ into the equation $\frac{r_a^2 + r_b^2}{T} = \frac{r_c^2}{T'}$,

$$r_c = \sqrt{r_a^2 + r_b^2}.$$

4. Balloon**

Let the horizontal length of the deformation be h. Then, the radius r of the circular cross-section which is in contact with the wall is

$$r = \sqrt{h(2R - h)}$$

by the intersecting chords theorem. The excess pressure in the balloon, relative to the atmosphere, is

$$\Delta p = \frac{2\gamma}{R}.$$

The force by the balloon on the wall is $\Delta p \pi r^2$. The force by the wall on the balloon is of the same magnitude but is opposite in direction. Thus, the equation of motion of the balloon is

$$m\ddot{h} = -\Delta p \pi r^2 = -\frac{2\gamma}{R}\pi h(2R - h).$$

Discarding the second-order term in h as h is small,

$$\ddot{h} = -\frac{4\pi\gamma}{m}h,$$

which is simple-harmonic. The time between the collision and rebound is half the period, and

$$\frac{T}{2} = \frac{\pi}{\omega} = \frac{\pi}{\sqrt{\frac{4\pi\gamma}{m}}} = \sqrt{\frac{\pi m}{4\gamma}}.$$

5. Capillary Forces between Parallel Plates**

When $d \ll D$, the shape of the liquid-air interface in a cross-section is a spherical cap of radius $R = \frac{d}{2\cos\theta}$. Therefore, the pressure within the water is

$$p = p_0 + \frac{2\gamma\cos\theta}{d},$$

by the Young-Laplace equation (the inverse of the other radius of curvature is negligible as compared to $\frac{1}{d}$ since $d \ll D$). The excess pressure hence exerts a force

$$F = (p - p_0) \cdot \frac{\pi D^2}{4} = \frac{\gamma\pi D^2\cos\theta}{2d},$$

on each of the plates, tending to bring them together. Note that the curvature pressure on a plate is zero due to its flatness. Actually, we should also include the adhesive forces which are associated with a force per unit length $\gamma\sin\theta$ along the three-phase interface, normal to each plate. This amounts to a total force $\gamma\sin\theta \cdot \pi D = \gamma\pi D\sin\theta$ on each plate in the normal direction — negligible when compared to F as $d \ll D$.

6. Accelerating Cylinder**

Define the positive x-axis to be directed leftwards and the origin to be at the right end of the cylindrical container. Consider an infinitesimal box element at coordinates (x, y, z). Considering the forces in the x-direction,

$$[p(x) - p(x + dx)]dA = \rho dA dx a$$
$$\implies \frac{dp}{dx} = -\rho a.$$

We have written p as a function of x only, as the pressure must be independent of y and z for the forces on a fluid element to be balanced along those directions. Next, we also know that at the base of the cylinder, $p = p_0 + \rho gh$ where h is the height of the vertical fluid column above the point of concern. Thus,

$$\frac{dp}{dx} = \rho g\frac{dh}{dx}$$
$$\frac{dh}{dx} = -\frac{a}{g}.$$

This is the slope of the surface of the cylindrical segment.[13] The length of segment $\overline{AB}$ is thus $\frac{2Ra}{g}$. To calculate $\overline{CD} = l'$, we first compute the volume of the top of the cylindrical segment (which is shaped like a wedge). Notice that two of these form a cylinder of height $\frac{2Ra}{g}$ and radius R. Thus, its volume is

$$\frac{\pi R^3 a}{g}.$$

The remaining volume is

$$\pi R^2 l - \frac{\pi R^3 a}{g}.$$

Therefore,

$$l' = l - \frac{Ra}{g}.$$

If $a > \frac{lg}{R}$, the fluid can no longer cover the entire base of the cylinder.

7. Ball in Hole**

Generally, the ball experiences an upthrust, a normal force due to the bottom of the tank and its weight. When the upthrust alone is enough to balance its weight, the ball rises out of the hole.

In order to compute the upthrust due to the immersed portion of the sphere, first remove the portion of the sphere that is not immersed as that is irrelevant. If the bottom circle, produced by this truncation, were to be covered with a continuous section of water, the upthrust on the sphere would be $\rho g V$ where V is the volume of the sphere immersed in water. However, instead of water of pressure $p_0 + \rho g h$ (where p_0 is the atmospheric pressure), the truncated sphere only experiences a force on its bottom circle due to the atmosphere which only has pressure p_0. Therefore, the total upthrust experienced by the original sphere must be

$$F = \rho g V - \rho g h \cdot \pi r^2,$$

as we have to deduct the contribution from the missing pressure. V can be computed via simple integration. Consider a circle of radius R centered at

[13] Alternatively, we could have derived this from the fact that the liquid surface must be perpendicular to the effective gravity.

the origin of the xy-plane.

$$x^2 + y^2 = R^2$$
$$\implies y = \pm\sqrt{R^2 - x^2}.$$

Observe that we can obtain the volume of the immersed portion of the sphere (a truncated sphere) by rotating the parts of this circle from $x = -\sqrt{R^2 - r^2}$ to $x = R$ about the x-axis for π radians. Therefore,

$$\begin{aligned} V &= \pi \int_{-\sqrt{R^2-r^2}}^{R} y^2 dx \\ &= \pi \int_{-\sqrt{R^2-r^2}}^{R} (R^2 - x^2) dx \\ &= \left[\pi \left(R^2 x - \frac{x^3}{3}\right)\right]_{-\sqrt{R^2-r^2}}^{R} \\ &= \frac{\pi}{3}\left[2R^3 + (2R^2 + r^2)\sqrt{R^2 - r^2}\right]. \end{aligned}$$

We require

$$mg = \rho g V - \rho g h \cdot \pi r^2$$
$$\implies h = \frac{V}{\pi r^2} - \frac{m}{\rho \pi r^2} = \frac{2R^3}{3r^2} + \frac{2R^2 + r^2}{3r^2}\sqrt{R^2 - r^2} - \frac{m}{\rho \pi r^2}.$$

8. Melting Ice*

Let m_i and m_o denote the mass of the ice and the object inside the ice, respectively. Then,

$$m_i + m_o = \rho_w V_{dis},$$

where ρ_w is the density of water and V_{dis} is the volume of fluid displaced. Rearranging,

$$V_{dis} = \frac{m_i}{\rho_w} + \frac{m_o}{\rho_w}.$$

In all cases, let the density of the object be ρ_o.

In the first case, the bead is denser and will thus fall to the bottom after the ice melts. Then, the volume "released" from the melting of the ice cube is

$$\frac{m_i}{\rho_w} + \frac{m_o}{\rho_o},$$

which is evidently smaller than V_{dis} as $\rho_o > \rho_w$. The first term is the volume of the water due to the ice melting and the second term is the volume of the bead. Therefore, the additional volume is unable to fill up the displaced volume — implying that the fluid level drops.

In the second case, $\rho_o < \rho_w$. Then, the object will float while displacing a new volume of water

$$V'_{dis} = \frac{m_o}{\rho_w}$$

after the ice has melted.

The volume released is that of water due to the ice, $\frac{m_i}{\rho_i}$. Since this exactly compensates for the change in the volume of water displaced, the fluid level remains the same.

In the last case, the oil will not displace any water, as it will float on top of it. Therefore, the volume released is

$$\frac{m_i}{\rho_w} + \frac{m_o}{\rho_o} > V_{dis},$$

as $\rho_o < \rho_w$. Therefore, the fluid level (which includes the oil) rises.

9. Rock in Bowl*

h_1 comes from the volume displaced by the bowl-cum-rock system.

$$h_1 = \frac{m_r + m_b}{\rho_w A}.$$

Meanwhile, h_2 originates from the volume of the rock and the volume displaced by the bowl.

$$h_2 = \frac{m_r}{\rho_r A} + \frac{m_b}{\rho_w A}$$

$$h_1 - h_2 = \frac{m_r}{A}\left(\frac{1}{\rho_w} - \frac{1}{\rho_r}\right).$$

10. Moving Ants*

The total mass of the ants can be calculated from Archimedes' Principle.

$$m = \frac{\sqrt{3}}{4}\rho l^2 w.$$

For the wedge to be in stable equilibrium, the center of mass of the system must be lower than the center of buoyancy so that the torque due to the upthrust after a slight rotation of the wedge can help to correct the deviation of the wedge from the vertical. The center of buoyancy is located at the center of mass of the displaced fluid which is two-thirds of the vertical height, away from the bottom vertex. In the optimal configuration, suppose that mass m_1 is at the top edge. Then, $m - m_1$ mass must be at the bottom vertex for the configuration to be optimal. In the boundary case where the center of buoyancy and the center of mass of the system coincide,

$$m_1 = \frac{2}{3}m.$$

Thus, the maximum is

$$m_1 = \frac{\sqrt{3}}{6}\rho l^2 w.$$

11. Oscillation between Two Fluids**

Define the y-axis to be the vertical axis and define the origin to be at the interface between the two fluids. Now, define y as the coordinate of the bottom of the cylinder (immersed in the bottom fluid), positive below the origin. Then, the length of the cylinder immersed in the top fluid is $l - y$. The equation of motion of the cylinder is

$$m\ddot{y} = mg - \rho_1(l - y)Ag - \rho_2 yAg$$

$$\ddot{y} = \left(1 - \frac{\rho_1 Al}{m}\right)g - \frac{(\rho_2 - \rho_1)Ag}{m}y.$$

Using the substitution $u = y - \frac{m-\rho_1 Al}{A(\rho_2-\rho_1)}$,

$$\ddot{u} = -\frac{(\rho_2 - \rho_1)Ag}{m}u.$$

Thus if $\rho_2 > \rho_1$, the cylinder undergoes simple harmonic motion with angular frequency

$$\omega = \sqrt{\frac{(\rho_2 - \rho_1)Ag}{m}}$$

about an equilibrium $y = \frac{m-\rho_1 Al}{A(\rho_2-\rho_1)}$. Else, if $\rho_1 = \rho_2$, the cylinder is in a neutral equilibrium. Finally, if $\rho_2 < \rho_1$, the cylinder is in an unstable equilibrium.

12. Spraying Water*

Since the container is large, the velocity of the water in the container is essentially zero. Applying Bernoulli's equation to the points adjacent to the hole inside and outside of the container,

$$[p_0 + \rho g(H-h)] + \rho gh = p_0 + \rho gh + \frac{1}{2}\rho v^2.$$

The term in brackets on the left-hand side is the pressure inside the container that is adjacent to the hole while the second term is its gravitational potential energy term. Solving, the fluid velocity immediately outside the hole is

$$v = \sqrt{2g(H-h)}.$$

The ejected fluid then undergoes projectile motion. The time required for the water to reach the ground can be computed via the equation

$$h = \frac{1}{2}gt^2$$

$$t = \sqrt{\frac{2h}{g}}.$$

The range is then

$$R = vt = 2\sqrt{h(H-h)} = 2\sqrt{\frac{H^2}{4} - \left(h - \frac{H}{2}\right)^2}.$$

Evidently, the range attains its maximum value of $\frac{H}{2}$ when

$$h = \frac{H}{2}.$$

For the second part, the trajectory of a section of water that is released from the hole at height h is described by the equations

$$y = h - \frac{1}{2}gt^2,$$

$$x = vt = \sqrt{2g(h-h)}t,$$

where y is the height above the base of the container, x is the horizontal distance from the container and t is the time elapsed from the release of the water from the hole. We can rewrite the above as

$$y = h - \frac{x^2}{4(H-h)}$$
$$4h^2 - 4(y+H)h + 4Hy + x^2 = 0.$$

Suppose that we want to solve for h for a given point (x, y) along a trajectory. A solution exists for h only if the discriminant of the above is greater or equal to zero, hence

$$16(y+H)^2 - 64Hy - 16x^2 \geq 0$$
$$\implies (H-y)^2 \geq x^2.$$

Since $H > y$, the above requires

$$H - y \geq x.$$

Therefore, for a given value of y, the maximum horizontal distance of a point on a trajectory from the container is

$$x = H - y.$$

Thus, the envelope takes the form of a straight line delineated by the equation $y = H - x$.

13. Venturi-meter*

Applying Bernoulli's equation to the bottoms of the first and second tubes,

$$p_0 + \rho g h_1 + \frac{1}{2}\rho v_1^2 = p_0 + \rho g h_2 + \frac{1}{2}\rho v_2^2.$$

For steady flow,

$$A_1 v_1 = A_2 v_2 = Q,$$

where Q is the volume flow rate.

$$\rho g(h_2 - h_1) = \frac{1}{2}\rho Q^2 \left(\frac{1}{A_1^2} - \frac{1}{A_2^2}\right).$$

Then,

$$Q = \frac{A_1 A_2 \sqrt{2g(h_2 - h_1)}}{\sqrt{A_2^2 - A_1^2}}.$$

14. Pitot Tube*

Let the pressure of the incoming gas inside the vessel be P. At the tip of the pitot tube, the air is stationary. Thus, the pressure at this tip (inside the pitot tube) is

$$P' = P + \frac{1}{2}\rho_a v^2$$

by Bernoulli's principle. Furthermore, we can calculate P' by considering the difference in the heights of the fluid levels.

$$P' = P + \rho_l g \Delta h,$$

where the pressure difference due to the different heights of air columns in the tubes have been neglected. Comparing these expressions,

$$v = \sqrt{\frac{2\rho_l g \Delta h}{\rho_a}}.$$

15. Filling a Tank**

Let h_0 and A_t be the maximum height and cross-sectional area of the tank respectively. The volume rate of water flowing out of the tap is

$$W = \frac{h_0 A_t}{T_1}.$$

Now, let us analyze the second scenario where a hole of area A_h is made on the bottom of the tank. When the water level in the tank is at height $h(t)$, the flow speed of water gushing out of the hole is $\sqrt{2gh}$ by Bernoulli's principle. Therefore,

$$A_t \frac{dh}{dt} = -A_h \sqrt{2gh}.$$

Separating variables,

$$\int_{h_0}^{0} \frac{1}{\sqrt{h}} dh = \int_0^{T_2} -\frac{A_h}{A_t}\sqrt{2g}dt$$

$$2\sqrt{h_0} = \frac{A_h}{A_t}\sqrt{2g}T_2$$

$$T_2^2 = \frac{2h_0 A_t^2}{g A_h^2}.$$

Proceeding with the final set-up, the differential equation describing the instantaneous height in the tank is

$$A_t \frac{dh}{dt} = W - A_h \sqrt{2gh}.$$

Observe that the height in the tank will only stop increasing at

$$h = \frac{W^2}{2gA_h^2} = \frac{A_t^2 h_0^2}{2gA_h^2 T_1^2}.$$

This must be larger than h_0 for the tank to overflow, so

$$\frac{A_t^2 h_0^2}{2gA_h^2 T_1^2} > h_0$$

$$\frac{T_2^2}{4T_1^2} > 1$$

$$\frac{T_1}{T_2} < \frac{1}{2}.$$

16. Sucking Water**

The water inside the tube is essentially stationary relative to the tube, as the aperture is small. Let the pressure at the top end of the tube be p_2 (this is near the hole but still inside the tube). Consider an infinitesimal volume element that is between radial distances r and $r + dr$ from the central portion of the tube. Then, it experiences pressures p and $p + dp$ on its faces, both of area dA. The net force due to pressure must produce the required centripetal force. Thus,

$$dp dA = \rho dr dA r \omega^2$$

$$\int_{p_1}^{p_2} dp = \int_0^l \rho r \omega^2 dr$$

$$p_2 = p_1 + \frac{\rho l^2 \omega^2}{2},$$

where p_1 is the pressure in the bend of the tube. Furthermore,

$$p_1 = p_0 - \rho g h,$$

where p_0 is the atmospheric pressure. Thus,

$$p_2 = p_0 - \rho g h + \frac{\rho l^2 \omega^2}{2}.$$

Now, consider the situation in the rotating frame. By applying Bernoulli's principle to the inside of the top end of the tube and immediately outside the hole,

$$p_2 = p_0 + \frac{1}{2}\rho l^2 v^2,$$

where v is the emitted velocity in the rotating frame.

$$v = \sqrt{l^2\omega^2 - 2gh}.$$

The velocity in the lab frame is thus

$$v_{lab} = \sqrt{v^2 + l^2\omega^2} = \sqrt{2l^2\omega^2 - 2gh},$$

as there is an additional tangential velocity $l\omega$ due to the rotation of the tube. To compute the power required, we first calculate the volume flow rate which is

$$Q = Av = A\sqrt{l^2\omega^2 - 2gh}.$$

Note that this is not Av_{lab}. In a time interval dt, Qdt amount of fluid moves through the entire tube. The external power is only responsible for the additional kinetic energy of the water associated with its tangential velocity due to the rotation. This additional kinetic energy is essentially that of the ejected fluid.

$$P = \frac{1}{2}\rho Q l^2\omega^2 = \frac{1}{2}\rho A\sqrt{l^2\omega^2 - 2gh}\,l^2\omega^2.$$

17. Transmitting Water**

Apply Bernoulli's equation to the fluid level of the first container and the immediate vicinity of the hole in the second container. Since the cross-sectional area of the first container is large as compared to the hole, the flow velocity in the first container can be neglected.

$$p_0 + \rho g h_1 = p_0 + \rho g h_2 + \frac{1}{2}\rho v^2,$$

where h_1 and h_2 are the heights of the water level in the respective containers and v is the flow velocity into the second container through the hole.

$$v = \sqrt{2g(h_1 - h_2)}.$$

The volume flux is

$$Av = A\sqrt{2g(h_1 - h_2)} = -A_1\frac{dh_1}{dt} = A_2\frac{dh_2}{dt},$$

as the total volume must be conserved. Then,

$$A_1A_2\frac{d(h_2-h_1)}{dt} = (A_1+A_2)Av = (A_1+A_2)A\sqrt{2g(h_1-h_2)}.$$

Using the substitution $x = h_1 - h_2$,

$$\dot{x} = -\frac{(A_1+A_2)A}{A_1A_2}\sqrt{2gx}.$$

Separating variables and integrating,

$$\int_{x_0}^{0}\frac{1}{\sqrt{x}}dx = -\int_0^t \frac{(A_1+A_2)A}{A_1A_2}\sqrt{2g}dt$$

$$t = \sqrt{\frac{2x_0}{g}}\frac{A_1A_2}{(A_1+A_2)A}.$$

18. Curving Water**

Let the velocities of the fluid at the left and right ends be $\boldsymbol{u}$ and $\boldsymbol{v}$ respectively. The net force on the portion of fluid is the rate of change of momentum. Let Q be the volume flow rate. In time dt, the net change in momentum of the entire portion of water is $\rho Q dt(\boldsymbol{v}-\boldsymbol{u})$. Essentially, the velocity of $\rho Q dt$ mass of water changes from $\boldsymbol{u}$ to $\boldsymbol{v}$. If the x and y-axes are defined to be positive rightwards and upwards respectively, the components of the net force in the corresponding directions are

$$F_{tot,x} = \rho Q(v\cos\theta - u),$$
$$F_{tot,y} = \rho Q v\sin\theta.$$

Since $Q = A_1u = A_2v$,

$$F_{tot,x} = \rho Q^2\left(\frac{1}{A_2}\cos\theta - \frac{1}{A_1}\right),$$
$$F_{tot,y} = \frac{\rho Q^2\sin\theta}{A_2}.$$

Furthermore, from Bernoulli's equation,

$$p_1 + \frac{1}{2}\rho u^2 = p_2 + \frac{1}{2}\rho v^2,$$

where the terms regarding the heights of the fluid levels have been canceled since the pipe lies on a horizontal table.

$$\rho Q^2 = \frac{2(p_1-p_2)A_1^2A_2^2}{A_1^2-A_2^2}.$$

Substituting this into the equations for the net force components,

$$F_{tot,x} = \frac{2(p_1 - p_2)A_1A_2(A_1\cos\theta - A_2)}{A_1^2 - A_2^2},$$
$$F_{tot,y} = \frac{2(p_1 - p_2)A_1^2A_2\sin\theta}{A_1^2 - A_2^2}.$$

The force on the fluid due to the walls is the net force subtracted by the force due to the pressure at the two ends.

$$F_{wall,x} = F_{tot,x} - p_1A_1 + p_2A_2\cos\theta = \frac{p_1(2A_1^2A_2\cos\theta - A_1A_2^2 - A_1^3)}{A_1^2 - A_2^2}$$
$$- \frac{p_2(A_1^2A_2\cos\theta - 2A_1A_2^2 + A_2^3\cos\theta)}{A_1^2 - A_2^2},$$
$$F_{wall,y} = F_{tot,y} + p_2A_2\sin\theta = \frac{A_2\sin\theta(2p_1A_1^2 - p_2A_1^2 - p_2A_2^2)}{A_1^2 - A_2^2}.$$

19. Hydraulic Jump**

We first calculate the force due to pressure at both ends of the flow. We first consider the left end. The pressure at a height h below the surface level is ρgh. Thus, the total force on the left end is

$$F_{left} = \int_0^{h_1} \rho ghw dh = \frac{1}{2}\rho gwh_1^2.$$

Similarly, the force on the right end is

$$F_{right} = -\frac{1}{2}\rho gwh_2^2.$$

Next, we compute the change in momentum of the section of fluid between the two ends after a time interval dt. Let the volume flow rate be

$$Q = h_1wv_1 = h_2wv_2,$$

where v_2 is the average flow velocity at the right end. In time dt, the velocity of ρQdt mass of fluid effectively changes from v_1 to v_2, on average. Thus, the change in momentum in time dt is

$$dp = \rho Q dt(v_2 - v_1) = \frac{\rho Q^2}{w}\left(\frac{1}{h_2} - \frac{1}{h_1}\right)dt = \frac{\rho Q^2(h_1 - h_2)}{wh_1h_2}dt.$$

By the impulse-momentum theorem,

$$(F_{left} + F_{right})dt = dp$$

$$\implies \frac{1}{2}\rho gw(h_1^2 - h_2^2) = \frac{\rho Q^2(h_1 - h_2)}{wh_1h_2}$$

$$(h_1 - h_2)\left(h_1 + h_2 - \frac{2Q^2}{gw^2h_1h_2}\right) = 0.$$

Discarding the trivial solution $h_2 = h_1$ and rearranging the expression in the brackets on the right,

$$h_2^2 + h_1h_2 - \frac{2Q^2}{gw^2h_1} = 0.$$

Solving this quadratic equation and substituting $Q = h_1wv_1$ yields

$$h_2 = -\frac{h_1}{2} + \sqrt{\frac{h_1^2}{4} + \frac{2h_1v_1^2}{g}},$$

where we have rejected the negative solution which is physically incorrect.

Chapter 10

Oscillations

10.1 Simple Harmonic Motion

Simple harmonic motion is the periodic motion of an object whose acceleration along a certain direction, $\ddot{x}$, is proportional to its displacement from an equilibrium position in magnitude and opposite in direction to its displacement, x. We write the constant of proportionality as $-\omega^2$ for the sake of convenience and get

$$\ddot{x} = -\omega^2 x.$$

The general solution to this differential equation is

$$x = A\sin(\omega t + \phi), \tag{10.1}$$

where $A \geq 0$ is the amplitude of oscillation and ϕ is the initial phase angle or offset. Both constants are determined by the initial conditions on displacement and velocity. ω is termed the angular frequency of oscillation, which is a characteristic of the oscillating system and is independent of the initial conditions.

Proof: Using the common trick that

$$\ddot{x} = \frac{d\dot{x}}{dt} = \frac{d\dot{x}}{dx} \cdot \frac{dx}{dt} = \dot{x}\frac{d\dot{x}}{dx},$$

we get

$$\int \dot{x} d\dot{x} = \int -\omega^2 x dx$$

$$\frac{\dot{x}^2}{2} = -\frac{\omega^2 x^2}{2} + \frac{c^2}{2}$$

$$\frac{dx}{dt} = \pm\sqrt{c^2 - \omega^2 x^2}$$

$$\int \frac{1}{\sqrt{c^2 - \omega^2 x^2}} dx = \pm \int dt.$$

Using the substitution $x = \frac{c}{\omega}\sin\theta$, $dx = \frac{c}{\omega}\cos\theta d\theta$,

$$\int \frac{1}{c\cos\theta} \cdot \frac{c}{\omega}\cos\theta d\theta = \pm \int dt$$

$$\frac{\theta}{\omega} = \pm t + k$$

for some constant k. Substituting $\theta = \sin^{-1}\frac{\omega x}{c}$,

$$\sin^{-1}\frac{\omega x}{c} = \pm\omega t + \omega k$$

$$x = \frac{c}{\omega}\sin(\pm\omega t + \omega k) \implies x = A\sin(\omega t + \phi),$$

where the $\pm$ sign in front of the variable t has been absorbed into the initial phase angle ϕ as $\sin(-\omega t + \omega k) = \sin(\pi + \omega t - \omega k) = \sin(\omega t + \phi)$ where $\phi = \pi - \omega k$. Let us examine some of the terms in the above equation.

- A is the amplitude of the oscillation. It is the maximum magnitude of the displacement of an oscillating particle from an equilibrium position.
- ω is the angular frequency of oscillation which is the rate of change of the phase angle of oscillation. The period of an oscillation, T, refers to the time needed for one complete cycle while the frequency of an oscillation, f, refers to the number of complete cycles of oscillations per unit time. The angular frequency ω is related to these quantities in the following manner:

$$\omega = 2\pi f = \frac{2\pi}{T}. \tag{10.2}$$

- ϕ is the initial phase angle or phase offset. It is determined by the initial displacement and velocity. It gives a sense of where the oscillating particle is when $t = 0$ or when an observer starts his or her timer.
- The equilibrium point is the position where the object experiences no net acceleration or force. This occurs when $x = 0$.

10.1.1 *Relationships between Kinematic Quantities*

In certain situations, we may be interested in other quantities describing simple harmonic motion, such as the instantaneous velocity v and acceleration

a of the object. We can solve for them by taking the time derivatives of its displacement.

$$v = \omega A \cos(\omega t + \phi), \tag{10.3}$$

$$a = -\omega^2 A \sin(\omega t + \phi). \tag{10.4}$$

We see that there is a $\frac{\pi}{2}$ phase difference between the instantaneous velocity and the displacement of the object and a π phase difference between the instantaneous acceleration and the displacement of the object. If we let the initial phase angle be zero and plot the corresponding displacement, velocity and acceleration of the oscillating object against time, we obtain the following graphs (Fig. 10.1).

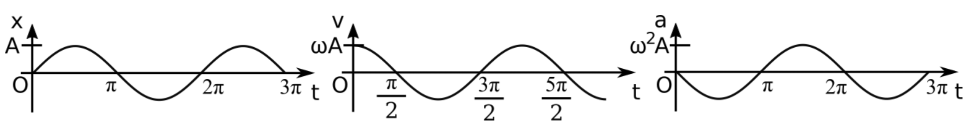


Figure 10.1: x, v and a against t graphs

Next, we might be interested in expressing the instantaneous velocity and acceleration of the object as functions of displacement instead. This is often more edifying as we can only physically observe the displacement of an object in a set-up most of the time.

$$v = \omega A \cos(\omega t + \phi) = \pm \omega A \sqrt{1 - \sin^2(\omega t + \phi)} = \pm \omega \sqrt{A^2 - x^2}, \tag{10.5}$$

$$a = -\omega^2 x.$$

If we plot the instantaneous velocity and acceleration of the oscillating body against its displacement, we see that we obtain an ellipse and a straight line, respectively (Fig. 10.2).

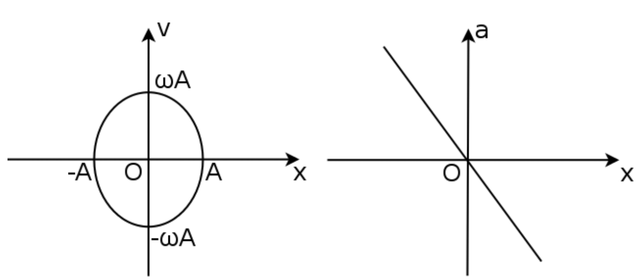


Figure 10.2: v and a against x graphs

To show that $v(x)$ is an ellipse, we can square Eq. (10.5) to obtain

$$v^2 + \omega^2 x^2 = \omega^2 A^2$$
$$\implies \left(\frac{v}{\omega A}\right)^2 + \left(\frac{x}{A}\right)^2 = 1^2,$$

which delineates an ellipse with axes length ωA and A along the v and x directions respectively.

Problem: Determine the possible displacements of the oscillating particle from the equilibrium position when its instantaneous speed is half of its maximum speed.

Since the maximum speed of the particle is $A\omega$, its instantaneous velocity is

$$v = \pm\frac{1}{2}A\omega$$
$$\pm\frac{1}{2}A\omega = \pm\omega\sqrt{A^2 - x^2}.$$

Evidently, we can only match the expressions of the same signs together. Then,

$$x^2 = \frac{3}{4}A^2$$
$$x = \pm\frac{\sqrt{3}}{2}A.$$

10.1.2 *Conservation of Energy*

With regard to the dynamics of simple harmonic motion, the total mechanical energy of a body undergoing simple harmonic motion is conserved. This is because the simple harmonic differential equation implies that the force, which acts on the oscillating body, is conservative. Recall that in our derivation of the general solution to the simple harmonic differential equation,

$$\frac{\dot{x}^2}{2} + \frac{\omega^2 x^2}{2} = \frac{c^2}{2},$$

which is a constant. Observe that the left-hand side is akin to the total mechanical energy of the particle divided by an inertial term: $\frac{\dot{x}^2}{2}$ is akin to a specific kinetic energy while $\frac{\omega^2 x^2}{2}$ is akin to a specific potential energy.

10.1.3 *Examples of Simple Harmonic Set-ups*

Now that we have studied the kinematics of simple harmonic motion, let us look at some realistic situations where this ubiquitous periodic motion arises.

A standard procedure in deriving the simple harmonic differential equation would be to first write down the equation of motion of the system with respect to a particular coordinate, which we shall denote as x for now. Then, an x-coordinate x_0 that corresponds to a state of stable equilibrium is identified. For systems whose regime of simple harmonic motion is the immediate vicinity of the equilibrium position, a small deviation of the system from x_0 is considered, such that the x-coordinate of the system can be represented as $x = x_0 + \varepsilon$, where ε is a small deviation. Such systems are known to exhibit small oscillations. Maclaurin expansions are then performed while discarding second-order terms in ε to generate the required simple harmonic differential equation. In certain systems where simple harmonic motion is exact, we can consider a general displacement of the particle from its equilibrium position to reach the simple harmonic differential equation.

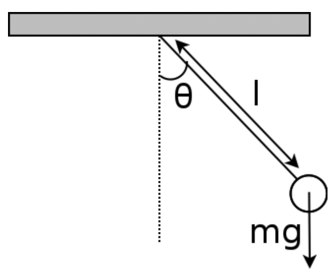


Figure 10.3: Simple pendulum

Consider the classic pendulum — a ball of negligible size and mass m is attached to a massless string of length l that is attached to the ceiling. We can show that at small angular displacements from the vertical (which is a stable equilibrium position), this pendulum exhibits simple harmonic motion.

Analyzing forces in the tangential direction,

$$m(r\ddot{\theta} + 2\dot{r}\dot{\theta}) = -mg\sin\theta.$$

Since the length of the string remains constant, $r = l$.

$$ml\ddot{\theta} = -mg\sin\theta.$$

For small angles, we can use the Maclaurin series to approximate $\sin\theta \approx \theta$ such that

$$ml\ddot{\theta} = -mg\theta \implies \ddot{\theta} = -\frac{g}{l}\theta.$$

This is the simple harmonic differential equation. Solving it gives

$$\theta = A \sin\left(\sqrt{\frac{g}{l}}t + \phi\right).$$

Thus, we see that the angular frequency of a simple pendulum is

$$\omega = \sqrt{\frac{g}{l}},$$

which is in fact independent of the mass of the point mass attached to the string.

In the general case of an object attached to a pivot, the equation $\tau = I\ddot{\theta}$ is more useful in determining the angular frequency of oscillations and other related quantities. Both the moment of inertia I and the torque τ on the system should be computed with respect to the fixed pivot.

Problem: A uniform rod of mass m and length l is attached to a pivot at one of its ends. Determine the angular frequency of small oscillations of the rod about the vertical orientation.

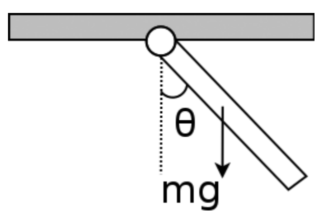


Figure 10.4: Rod pendulum

Recall that the moment of inertia of a uniform rod about its center is $\frac{1}{12}ml^2$. Therefore, the moment of inertia of the rod about one of its ends is

$$I = \frac{1}{12}ml^2 + m\frac{l^2}{4} = \frac{1}{3}ml^2$$

by the parallel axis theorem. The net torque on the system about the fixed pivot is due to that of the weight of the rod, acting at its center of mass. Therefore, the net torque on the rod is

$$\tau = -\frac{1}{2}mgl\sin\theta.$$

Applying the equation $\tau = I\ddot{\theta}$ (as the pivot is an ICoR) and the small angle approximation $\sin\theta \approx \theta$,

$$\ddot{\theta} = -\frac{3g}{2l}\theta.$$

The angular frequency is then

$$\omega = \sqrt{\frac{3g}{2l}}.$$

Such systems involving extended dangling objects are known as physical pendulums. It is not difficult to generalize the above result to show that the angular frequency of a physical pendulum is

$$\omega = \sqrt{\frac{Mr_{CM}g}{I_{pivot}}}, \tag{10.6}$$

where M is the total mass of the object, r_{CM} is the distance between the pivot and the center of mass of the object and I_{pivot} is the moment of inertia of the object about the fixed pivot.

Two-Dimensional Systems

The small oscillations of inherently two-dimensional systems whose equations of motion consist of two variables are usually considered in the context of perturbing a single coordinate. Then, variables in terms of the other coordinate should be eliminated to obtain a differential equation in terms of the single, relevant coordinate. This elimination is usually performed via the conservation of angular momentum in a certain direction.

Problem: A planet of mass m is currently undergoing circular motion about the Sun of mass M at a radius r_0. If the planet is somehow given a slight radial displacement, determine its angular frequency of small oscillations in the radial direction. What is the resultant trajectory of the planet?

The radial equation of motion of the planet in polar coordinates is

$$-\frac{GMm}{r^2} = m(\ddot{r} - r\dot{\theta}^2).$$

Since the push is only radial and because the gravitational force is central, the angular momentum of the planet relative to the Sun remains constant.

$$L = mr^2\dot{\theta}.$$

The radial equation of motion becomes

$$-\frac{GM}{r^2} = \ddot{r} - \frac{L^2}{m^2r^3}.$$

Furthermore, we know that when $r = r_0$, $\ddot{r} = 0$. Then,

$$\frac{GM}{r_0^2} = \frac{L^2}{m^2r_0^3}.$$

This will be useful in canceling some terms later — a common denominator in all small oscillation problems. Suppose that the radius now becomes[1] $r_0 + \varepsilon$. The equation of motion becomes

$$-\frac{GM}{r_0^2\left(1+\frac{\varepsilon}{r_0}\right)^2} = \ddot{\varepsilon} - \frac{L^2}{m^2r_0^3\left(1+\frac{\varepsilon}{r_0}\right)^3}.$$

Performing a Maclaurin expansion and discarding second-order terms in $\frac{\varepsilon}{r_0}$,

$$-\frac{GM}{r_0^2} + \frac{2GM}{r_0^3}\varepsilon = \ddot{\varepsilon} - \frac{L^2}{m^2r_0^3} + \frac{3L^2\varepsilon}{m^2r_0^4}.$$

Observe that the first term on the left-hand side cancels the second term on the right-hand side by our previous equation for r_0. Then,

$$\begin{aligned}\ddot{\varepsilon} &= \left(\frac{2GM}{r_0^3} - \frac{3L^2}{m^2r_0^4}\right)\varepsilon \\ &= \left(\frac{2GM}{r_0^3} - \frac{3GM}{r_0^3}\right)\varepsilon \\ &= -\frac{GM}{r_0^3}\varepsilon.\end{aligned}$$

Thus, the angular frequency of small oscillations in the radial direction is

$$\omega = \sqrt{\frac{GM}{r_0^3}}.$$

At first glance, one might expect the resultant trajectory of the planet to take the form of a "flower pattern" as the planet oscillates radially along an originally circular orbit. However, observe that the period of the original circular orbit is exactly ω! Therefore, the planet only attains the maximum and minimum radial distance from the Sun once per complete revolution — indicating that the new orbit is an ellipse with semi-major and semi-minor axes $r_0 + |\varepsilon|$ and $r_0 - |\varepsilon|$! When $|\varepsilon| \ll r_0$, the eccentricity of the ellipse is virtually zero such that the new orbit is akin to a circle, with the Sun slightly displaced from the center of the circle.

This is intuitive from the perspective of Kepler's first law as the orbit of a planet is in general an ellipse, with the Sun as a focus. Note that the resultant orbit of the planet must still be bounded as a slight deviation

[1] Note that in the previous set-ups, we did not have to explicitly state this, as the equilibrium angles were $\theta = 0$.

imparts negligible mechanical energy to it. In retrospect, we could have also imposed the condition that the trajectory can only be an ellipse, to conclude that the angular frequency of radial oscillations must match the angular frequency of the planet's original orbit!

10.2 Deriving Angular Frequency from Potential Energy

When a particle is solely under the influence of conservative forces, we can associate it with a potential energy function that is strictly only a function of its position. Let us consider an arbitrary potential energy function $U(x)$ for a one-dimensional system while keeping in mind that the net force on the particle is $F = -\frac{dU}{dx}$. The equilibrium positions correspond to the stationary points of the $U(x)$ graph as $F = -\frac{dU}{dx} = 0$ there.

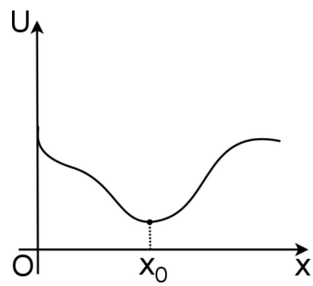


Figure 10.5: One-dimensional potential energy

In light of our goal of analyzing oscillations, let us observe the motion of the particle at positions in the vicinity of a minimum as the particle strives towards attaining a lower potential energy — implying that the minimum corresponds to a stable equilibrium state as any deviation from the minimum tends to be minimized. Assuming that the potential energy function has a minimum at x_0, we can expand $U(x)$ as a Taylor series about x_0.

$$U(x) = U(x_0) + U'(x_0)(x - x_0) + \frac{U''(x_0)}{2}(x - x_0)^2 + \cdots$$

where we will neglect third-order terms as we assume that the particle is near x_0. Now let us consider the force on the particle in the x-direction.

$$F = -\frac{dU}{dx} = -U'(x_0) - U''(x_0)(x - x_0).$$

Since x_0 is a minimum, $U'(x_0) = 0$. Furthermore, if we use a change of variables $\varepsilon = x - x_0$ such that $F = m\ddot{x} = m\ddot{\varepsilon}$, we can simplify the equation above to

$$m\ddot{\varepsilon} = -U''(x_0)\varepsilon.$$

Evidently, this describes a simple harmonic motion with an angular frequency given by

$$\omega = \sqrt{\frac{U''(x_0)}{m}}. \tag{10.7}$$

Perhaps an intuitive and lucid explanation of why this should be the case is that if you zoom closer to the minimum, the regions around it will look like a parabola. Thus, the potential energy curve is approximately $U = A(x - x_0)^2 + c$ which gives $F = -2A(x - x_0)$ — a simple harmonic force (like a spring).

Problem: As its name implies, a spring-mass system consists of a mass m connected by a massless spring of spring constant k to a fixed pivot. Determine the angular frequency of oscillations if the system lies on a horizontal table and if it hangs vertically.

In both cases, define the origin at the equilibrium position such that x denotes the displacement of m in the relevant direction, from the equilibrium position. In the horizontal case, the potential energy at a displacement x is

$$U = \frac{1}{2}kx^2$$
$$U''(0) = k.$$

Therefore,

$$\omega = \sqrt{\frac{k}{m}}.$$

In the vertical case, the equilibrium position of the mass is $\frac{mg}{k}$ below the relaxed length of the spring. At a displacement x below this equilibrium position, the potential energy associated with the particle is

$$U = \frac{1}{2}k\left(\frac{mg}{k} + x\right)^2 - mgx$$
$$U''(0) = k,$$
$$\omega = \sqrt{\frac{k}{m}}.$$

Effective Potential for Central Force Systems

In the case of central force systems which are two-dimensional, the one-dimensional method above can be applied to radial oscillations through the introduction of an effective potential. In a central force system, the total mechanical energy is

$$\frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\theta}^2 + U(r) = E,$$

where $U(r)$ is the potential energy associated with the central force field and r is the radial position of the particle relative to the source of the central force field. Next, the angular momentum of the particle relative to the source is conserved.

$$mr^2\dot{\theta} = L.$$

Therefore, the first equation can be expressed as

$$\frac{1}{2}m\dot{r}^2 + \frac{L^2}{2mr^2} + U(r) = E,$$

which is akin to a one-dimensional conservation of energy equation in r with an effective potential

$$U_{eff}(r) = \frac{L^2}{2mr^2} + U(r).$$

Then, the above results can be directly applied to conclude that the angular frequency of radial oscillations (if the particle actually oscillates) is

$$\omega = \sqrt{\frac{U''_{eff}(r_0)}{m}},$$

where r_0 is the equilibrium radial coordinate.

Problem: Redo the previous problem on a planet orbiting the Sun using the effective potential method.

The effective potential is

$$U_{eff} = \frac{L^2}{2m^2r^2} - \frac{GMm}{r}.$$

Then,

$$U'_{eff} = -\frac{L^2}{m^2r^3} + \frac{GMm}{r^2}.$$

We know that at the equilibrium radial coordinate r_0, $U'_{eff} = 0$. Therefore,

$$\frac{L^2}{m^2 r_0^3} = \frac{GMm}{r_0^2}.$$

Moving on, we calculate the second derivative of U_{eff}.

$$U''_{eff} = \frac{3L^2}{m^2 r^4} - \frac{2GMm}{r^3}$$

$$U''_{eff}(r_0) = \frac{1}{r_0}\left(\frac{3L^2}{m^2 r_0^3} - \frac{2GMm}{r_0^2}\right) = \frac{GMm}{r_0^3}.$$

Therefore, the angular frequency of radial oscillations is

$$\omega = \sqrt{\frac{U''_{eff}(r_0)}{m}} = \sqrt{\frac{GM}{r_0^3}}.$$

10.3 Damped Oscillations

In an ideal oscillatory system, there are no dissipative forces and the total mechanical energy of the system is conserved. The system then oscillates indefinitely with a constant amplitude. However, in real-world systems, there are often dissipative forces which cause the amplitude of oscillation to gradually decrease over time. The resultant oscillatory motion is known as damped oscillations.

10.3.1 *Linear Differential Equations*

Before we solve the equation for damped oscillations, it might be helpful to know some properties of linear differential equations. A nth degree linear differential equation (with respect to time) takes the form:

$$c_n \frac{d^n x}{dt^n} + c_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + c_1 \frac{dx}{dt} + c_0 x = g(t),$$

where $g(t)$ may be a constant or a function of t. If $g = 0$, the equation is called a homogeneous equation. We shall first consider a homogeneous linear differential equation

$$c_n \frac{d^n x}{dt^n} + c_{n-1}\frac{d^{n-1}x}{dt^{n-1}} + \cdots + c_1 \frac{dx}{dt} + c_0 x = 0.$$

It turns out that the previous method in solving the simple harmonic differential equation is not applicable to the general case of linear differential equations. However, we can invoke a general theorem of linear differential equations which states that a nth degree homogeneous linear differential equation has exactly n linearly independent solutions. The next step is to determine these n solutions, by hook or by crook. The most general method is in fact guessing solutions of the form $x = e^{at}$. Then,

$$c_n a^n e^{at} + c_{n-1} a^{n-1} e^{at} + \cdots + c_1 a e^{at} + c_0 e^{at} = 0.$$

Dividing the equation by e^{at} throughout and invoking the fundamental theorem of algebra, this nth degree polynomial equation can be factorized into

$$(a - b_1)(a - b_2) \ldots (a - b_n) = 0,$$

which is also known as the characteristic equation. Thus, we have n roots for a which may be real, complex or even repeated. We can substitute any of these roots into the trial solution for x ($x = e^{at}$) and it would satisfy the differential equation. For example, $x = e^{b_1 t}$ and $x = e^{b_2 t}$ are solutions to the linear differential equation. Due to the linear nature of the differential equation, any linear combination of these solutions is also a solution. Thus, a general solution would take the form

$$x = z_1 e^{b_1 t} + z_2 e^{b_2 t} + \cdots + z_n e^{b_n t},$$

where the z_i's are possibly complex constants. Since a nth degree homogeneous linear differential equation has n linearly independent solutions, we are done if there are no repeated roots and the expression for x above is the most general solution to the homogeneous linear differential equation. However, when there are repeated roots, we have to search for other linearly independent solutions. Usually, if a root b_j is repeated k times, we guess the following k solutions: $e^{b_j t}, te^{b_j t}, t^2 e^{b_j t}, \ldots, t^{k-1} e^{b_j t}$.

Let us apply this guessing technique to the simple harmonic differential equation

$$\ddot{x} + \omega^2 x = 0.$$

Substituting e^{at} into x,

$$a^2 e^{at} + \omega^2 e^{at} = 0$$
$$a = \pm i\omega,$$

where $i = \sqrt{-1}$. Therefore, the general solution for x is

$$x = z_1 e^{i\omega t} + z_2 e^{-i\omega t},$$

where z_1 and z_2 are arbitrary constants which may be complex. Since x represents the displacement which is a physical quantity, it must be real at all times. Therefore, z_1 and z_2 must be complex conjugates.

$$z_1 = z_2^*.$$

Then, we can represent z_1 and z_2 as

$$z_1 = \frac{A}{2}e^{i\phi},$$

$$z_2 = \frac{A}{2}e^{-i\phi},$$

where A and ϕ are arbitrary real constants. Correspondingly,

$$x = \frac{A}{2}(e^{i(\omega t+\phi)} + e^{-i(\omega t+\phi)}) = A\cos(\omega t + \phi)$$

by Euler's identity, $e^{i\theta} = i\sin\theta + \cos\theta$. Observe that since z_1 and z_2 are complex conjugates and because x must be real, we have effectively taken the real component of either of them (up to a constant factor). Therefore, whenever we have a solution for a real x in terms of pairs of the form $z_1e^{i\omega t} + z_2e^{-i\omega t}$, x can be simply expressed in terms of the real component of one term from each pair, such as $\mathrm{Re}(z_1e^{i\omega t})$.

This, in combination with the following final remark, can prove to be extremely useful. If all constants c_i in a homogeneous linear differential equation are real[2] and if y is a solution, y^* must also be a solution to the equation. This can be easily proven by taking the complex conjugate of the entire differential equation.

$$\left(c_n\frac{d^ny}{dt^n}\right)^* + \left(c_{n-1}\frac{d^{n-1}y}{dt^{n-1}}\right)^* + \cdots + \left(c_1\frac{dy}{dt}\right)^* + (c_0y)^* = 0^*.$$

Since $(z_1z_2)^* = z_1^* \cdot z_2^*$ for two arbitrary complex numbers z_1 and z_2 and the constants c_i are real,

$$c_n\left(\frac{d^ny}{dt^n}\right)^* + c_{n-1}\left(\frac{d^{n-1}y}{dt^{n-1}}\right)^* + \cdots + c_1\left(\frac{dy}{dt}\right)^* + c_0y^* = 0.$$

The order of differentiation and complex conjugation does not matter. Therefore,

$$c_n\frac{d^ny^*}{dt^n} + c_{n-1}\frac{d^{n-1}y^*}{dt^{n-1}} + \cdots + c_1\frac{dy^*}{dt} + c_0y^* = 0,$$

which proves that y^* is also a solution. In such cases, we can simply take the real component of y in writing the general solution for x — neglecting its complex conjugate y^*.

[2]This is usually the case in physical scenarios.

Lastly, we consider the non-homogeneous linear differential equation

$$c_n \frac{d^n x}{dt^n} + c_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \cdots + c_1 \frac{dx}{dt} + c_0 x = g(t).$$

The general solution to this differential equation involves a particular solution and a homogenous solution. Firstly, we find a solution, x_p, which satisfies this equation. This is known as the particular solution. Then, we can add to the particular solution the homogeneous solution, obtained by letting $g = 0$ in the equation above, to obtain the general solution to the non-homogeneous linear differential equation. This is due to the fact that adding the homogeneous solution results in an additional value of zero on the right-hand side of the above equation — leaving it unchanged. Lastly, note that the particular solution does not depend on the initial conditions. Instead, the initial conditions are still encoded in the homogeneous solution.

10.3.2 *Equation of Motion*

In most cases, the damping force on an oscillating body is proportional to its velocity and acts in the opposite direction. Thus, the differential equation for the displacement of the body takes the form

$$\ddot{x} + 2\gamma\dot{x} + \omega^2 x = 0.$$

Guessing a solution of the form $x = e^{at}$, we obtain the characteristic equation

$$a^2 + 2\gamma a + \omega^2 = 0.$$

Solving, we obtain two possibly repeated roots:

$$a = \frac{-2\gamma \pm \sqrt{4\gamma^2 - 4\omega^2}}{2} = -\gamma \pm \sqrt{\gamma^2 - \omega^2}.$$

Now we have to consider three cases — namely when the term inside the square root is negative, positive and zero.

10.3.3 *Light Damping*

When $\gamma < \omega$, the system experiences light damping or is underdamped. Thus, we can rewrite the two solutions for a as

$$a = -\gamma \pm i\sqrt{\omega^2 - \gamma^2}.$$

Then, our general solution for the equation of motion is

$$\begin{aligned} x &= c_1 e^{-\gamma t + i\sqrt{\omega^2-\gamma^2}t} + c_2 e^{-\gamma t - i\sqrt{\omega^2-\gamma^2}t} \\ &= e^{-\gamma t}\left(c_1 e^{i\sqrt{\omega^2-\gamma^2}t} + c_2 e^{-i\sqrt{\omega^2-\gamma^2}t}\right). \end{aligned}$$

For our displacement to be strictly real at all instances, c_1 and c_2 have to be complex conjugates. Letting $c_1 = \frac{c_0}{2}e^{i\phi}$ and $c_2 = \frac{c_0}{2}e^{-i\phi}$,

$$x = e^{-\gamma t}\frac{c_0}{2}\left(e^{i(\sqrt{\omega^2-\gamma^2}t+\phi)} + e^{-i(\sqrt{\omega^2-\gamma^2}t+\phi)}\right),$$

$$x = e^{-\gamma t}c_0 \cos\left(\sqrt{\omega^2-\gamma^2}t+\phi\right), \tag{10.8}$$

where we have invoked the elegant Euler's identity. Again, we have effectively taken the real part of one of the solutions. We see that the amplitude of oscillation decreases exponentially over time and that the angular frequency of oscillation is smaller than that of an undamped system. The angular frequency of the underdamped system ω_d is given by

$$\omega_d = \sqrt{\omega^2-\gamma^2}.$$

The displacement of the object as a function of time is illustrated below.

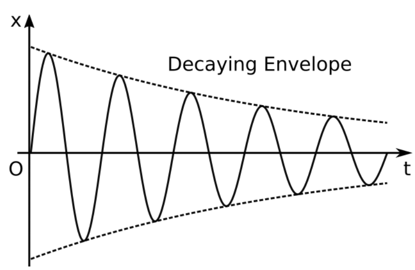


Figure 10.6: Light damping

The graph represents an oscillation bounded by an exponentially decreasing envelope which governs the amplitude of oscillation at every instant in time.

10.3.4 *Heavy Damping*

When $\gamma > \omega$, the system experiences heavy damping or is overdamped. There are two real solutions for a:

$$a = -\gamma \pm \sqrt{\gamma^2-\omega^2}.$$

Thus, our general solution to the equation of motion is

$$x = c_1 e^{-\left(\gamma - \sqrt{\gamma^2 - \omega^2}\right)t} + c_2 e^{-\left(\gamma + \sqrt{\gamma^2 - \omega^2}\right)t}, \tag{10.9}$$

where c_1 and c_2 are real constants. The system does not undergo oscillatory motion as the exponents are real. As time passes by, the displacement gradually tends to 0. This decay is indeed extremely gradual as the significant term in the long run is $e^{-\left(\gamma - \sqrt{\gamma^2 - \omega^2}\right)t}$. A vivid example of an overdamped system would be a door that takes an incredibly long time to close. On another note, it is intriguing to show that the oscillating particle can cross the origin at most once. Substituting $x = 0$ and simplifying,

$$\begin{aligned} c_1 e^{\sqrt{\gamma^2 - \omega^2}t} &= -c_2 e^{-\sqrt{\gamma^2 - \omega^2}t} \\ e^{2\sqrt{\gamma^2 - \omega^2}t} &= -\frac{c_2}{c_1}, \end{aligned}$$

which has one solution only if $\frac{c_2}{c_1} < 0$ (i.e. they are of opposite signs). In fact, the condition is much stricter — if the initial displacement is positive, $c_1 < 0$ and $c_2 > 0$ for the oscillating body to cross the origin. This is because, if the body really crossed the origin, the amplitude of its displacement must tend to 0^- after a long time (i.e. approach 0 from below the t-axis) as it cannot cross the origin again. In the long term, $e^{-(\gamma - \sqrt{\gamma^2 - \omega^2})t}$ is more significant than the other exponential term. Therefore, c_1 must be negative — implying that c_2 is positive. If the initial displacement is negative, the converse occurs. Plotting the graph of displacement against time for an overdamped system, we get the following possible curves.

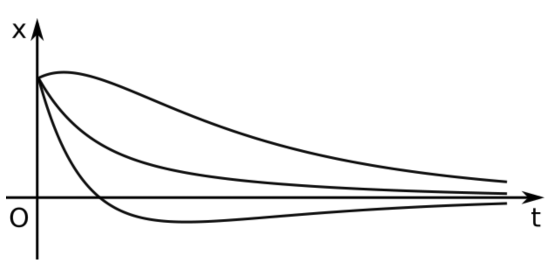


Figure 10.7: Heavy damping

10.3.5 *Critical Damping*

When $\gamma = \omega$, we say that the system is critically damped. a has only one solution $a = -\gamma$. Thus, we have only found one independent solution from the characteristic equation:

$$x = e^{-\gamma t}.$$

The other independent solution is $x = te^{-\gamma t}$ (you should verify this for yourself). Thus, the general solution for the equation of motion is

$$x = e^{-\gamma t}(c_1 + c_2 t), \tag{10.10}$$

where c_1 and c_2 are real constants. The exponential decay term in $c_2 te^{-\gamma t}$ dominates the $c_2 t$ term in the long run such that $x \to 0$ for very large t. The system does not oscillate at all and instead, returns to the equilibrium position in the shortest time possible, as compared to the other forms of damping. This can be proven by comparing the exponential decay constants. In the case of light damping, the exponential decay constant is γ in the regime $\gamma < \omega$ which is evidently smaller than $\gamma = \omega$ in the critical damping case. Furthermore, the dominant decay constant (the smaller one) in the case of overdamping is $\gamma - \sqrt{\gamma^2 - \omega^2}$ which can be shown to be smaller than ω.

$$\gamma - \omega < \gamma + \omega.$$

Since $\gamma > \omega$ in the regime of overdamping, we can multiply both sides by $\gamma - \omega > 0$ to obtain

$$\begin{aligned} (\gamma - \omega)^2 &< \gamma^2 - \omega^2 \\ \implies \gamma - \omega &< \sqrt{\gamma^2 - \omega^2} \\ \gamma - \sqrt{\gamma^2 - \omega^2} &< \omega. \end{aligned}$$

Thus, the particle returns to a state of equilibrium in the shortest time (still indefinitely long though) when the system is critically damped. Critical damping is paramount in many real systems, such as shock absorbers, in ensuring that a system stops immediately without oscillating about. On another note, the particle can, again, cross the origin at most once. When $x = 0$,

$$t = -\frac{c_1}{c_2},$$

which is only valid if $\frac{c_1}{c_2} < 0$. By a similar argument as above, if the initial displacement is positive, $c_1 > 0$ and $c_2 < 0$ for the oscillating particle to cross the origin, as the $e^{-\gamma t} c_2 t$ term is dominant over $e^{-\gamma t} c_1$ for large t. The possible displacement-against-time graphs of a critically damped system are depicted on the next page. They look roughly the same as those in the heavy damping case with the exception that the graph for small t is essentially linear. This is because, for small t, the exponential term in x is approximately unity such that $x \approx c_1 + c_2 t$.

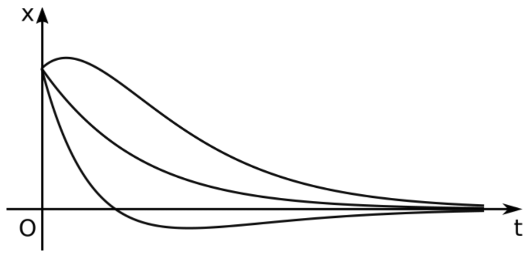


Figure 10.8: Critical damping

10.3.6 *Driven and Damped Oscillations*

Considering the fact that most real-world systems experience dissipative forces, an external periodic driving force is often applied to sustain their motion. Before solving the equation of motion for real driven oscillations, we shall first tackle a differential equation that takes the following form:

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = c_0 e^{i\omega_e t}.$$

We have added subscripts to the ω's to avoid confusion. Recapitulating, to solve a non-homogeneous linear differential equation, we need to determine a particular solution, before adding it to the homogeneous solution to procure the general solution. As there is a $e^{i\omega_e t}$ term on the right-hand side, it is wise to guess a particular solution $x_p = Ae^{i\omega_e t}$. In this case, we are solving for the constant A instead of the angular frequency. Substituting this trial solution into the expression, we get

$$-\omega_e^2 A + 2i\gamma\omega_e A + \omega_0^2 A = c_0 \implies A = \frac{c_0}{-\omega_e^2 + 2i\gamma\omega_e + \omega_0^2},$$

$$x_p = \frac{c_0}{-\omega_e^2 + 2i\gamma\omega_e + \omega_0^2} e^{i\omega_e t}.$$

Lastly, we can obtain the general solution to this differential equation by adding the appropriate homogeneous solution, derived previously, to the particular solution.

Now, we can consider providing a driving force to our oscillatory system of the form $F = f\cos\omega_e t$ where ω_e is known as the angular frequency of the external driving force. Then, the equation of motion of the oscillating body takes the form

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = c_0 \cos\omega_e t,$$

where c_0 is a real constant. ω_0 is known as the natural frequency of the system and is the angular frequency of oscillations when there are no damping and external driving forces. To solve this differential equation, consider the following auxiliary differential equation

$$\ddot{y} + 2\gamma\dot{y} + \omega_0^2 y = c_0 e^{i\omega_e t}.$$

Suppose that we have found a solution for y in the equation above. Then, we can take the real component of both sides.

$$\text{Re}(\ddot{y}) + \text{Re}(2\gamma\dot{y}) + \text{Re}(\omega_0^2 y) = \text{Re}\left(c_0 e^{i\omega_e t}\right).$$

Since the constants γ, ω_0 and c_0 are real and because the order of differentiating y and taking the real part of it does not matter,

$$\ddot{z} + 2\gamma\dot{z} + \omega_0^2 z = c_0 \cos\omega_e t,$$

where $z = \text{Re}(y)$ is the real component of y. Therefore, the solution for x that we desire is simply the real component of y! The particular solution for y was previously derived to be

$$y_p = \frac{c_0}{-\omega_e^2 + 2i\gamma\omega_e + \omega_0^2} e^{i\omega_e t}.$$

This can be expressed in a more suggestive form by applying Euler's formula to the denominator.

$$\begin{aligned} y_p &= \frac{c_0}{\sqrt{(\omega_0^2 - \omega_e^2)^2 + 4\gamma^2\omega^2 e^{i\phi}}} e^{i\omega_e t}, \\ &= \frac{c_0}{\sqrt{(\omega_0^2 - \omega_e^2)^2 + 4\gamma^2\omega^2}} e^{i(\omega_e t - \phi)}, \end{aligned}$$

where

$$\phi = \tan^{-1}\frac{2\gamma\omega_e}{\omega_0^2 - \omega_e^2}. \tag{10.11}$$

Therefore, the particular solution x_p can be obtained by taking the real component of the above.

$$x_p = \frac{c_0}{\sqrt{(\omega_0^2 - \omega_e^2)^2 + 4\gamma^2\omega_e^2}} \cos(\omega_e t - \phi). \tag{10.12}$$

In the long run, the damping terms will cause the homogeneous solution to tend to zero. Therefore given different initial conditions, the system will eventually reach the same steady state, which is described by its particular solution. Furthermore, in light of the decaying amplitude of the homogeneous solution, the particular solution is often valued over the homogeneous solution as it is a more enlightening description of the behaviour of driven oscillations — we shall therefore not concern ourselves too much with the general solution, which will tend to the particular solution in the long run.

Resonance

Consider a realistic oscillatory system. Because of damping, the amplitude of the system eventually diminishes to zero. Thus, we would like to sustain the motion of the system by applying a periodic force. However, what should the driving frequency ω_e (angular frequency of the periodic force) be, such that the resultant amplitude of oscillation is the greatest? This frequency is known as the resonant frequency. Resonance is a phenomenon in which an oscillatory system responds with a maximum amplitude to an external periodic force. The condition for resonance can be derived from Eq. (10.12). The amplitude of the particular solution of a driven damped oscillation, A, is

$$A = \frac{c_0}{\sqrt{(\omega_0^2 - \omega_e^2)^2 + 4\gamma^2\omega_e^2}} = \frac{c_0}{\sqrt{\left[\omega_e^2 - (\omega_0^2 - 2\gamma^2)\right]^2 + 4\omega_0^2\gamma^2 - 4\gamma^4}},$$

which attains the maximum value $\frac{c_0}{2\gamma\sqrt{\omega_0^2-\gamma^2}}$ when

$$\omega_e = \sqrt{\omega_0^2 - 2\gamma^2}.$$

Thus, $\sqrt{\omega_0^2 - 2\gamma^2}$ is the resonant frequency, ω_r. Observe that when $\gamma \ll \omega$ (no damping or light damping), $\omega_r \approx \omega_0$ is the condition for resonance. Thus, resonance occurs for underdamped and simple oscillations when the driving frequency is approximately equal to the natural frequency of the system.

When the system is underdamped and $\omega_e = \omega_r \approx \omega_0$, $\phi \to \frac{\pi}{2}$ as a result[3] of Eq. (10.11). The displacement of the oscillating body lags behind that of the driving force by a quarter of a cycle while the velocity of the body and the driving force are perfectly in phase. When the force is at its maximum, the displacement of the object is zero and it thus possesses the greatest velocity (in the same direction as the external driving force). This is intuitive from the standpoint of energy as the force should act with the greatest magnitude on the object when it is traveling the fastest, to maximize the work done by the driving force.

Lastly, we can plot the amplitude of driven oscillations, A, against the driving frequency ω_0 for different damping constants γ (Fig. 10.9).

As the system experiences greater damping, the amplitude decreases, the resonance peak becomes broader and the resonant frequency slightly

[3] $\tan\phi$ tends to $+\infty$ as $\omega_e \to \omega_0^-$. Note that ϕ must be in the first quadrant ($\frac{\pi}{2}$) and not the third ($\frac{3\pi}{2}$) as both the real and complex components of $e^{i\phi}$ are positive. Refer to the specific juncture at which we substituted $e^{i\phi}$ for further clarifications.

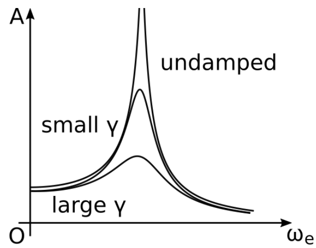


Figure 10.9: Maximum amplitude against driving frequency

decreases. In the case of undamped oscillations, the amplitude tends to infinity when $\omega_e \to \omega_0$ from both sides.

Summary

- A lightly damped system oscillates with an amplitude that is exponentially decaying. Its angular frequency is slightly smaller than the natural frequency of oscillation.
- An overdamped or heavily damped system gradually returns to the equilibrium position.
- A critically damped system returns to the equilibrium position in the least possible time.
- Resonance is a phenomenon where an oscillatory system responds with the greatest amplitude to an external driving force. The angular frequency of the driving force is known as the driving frequency. The angular frequency at which resonance occurs is known as the resonant frequency.
- The angular frequency of a damped oscillation ω_d and the resonant frequency ω_r of a damped system are

$$\omega_d = \sqrt{\omega_0^2 - \gamma^2},$$

$$\omega_r = \sqrt{\omega_0^2 - 2\gamma^2},$$

  where ω_0 is the natural frequency of the system refering to the angular frequency of the oscillation if there were no damping or driving forces.

10.4 Coupled Oscillations

In certain cases, we may have periodic systems which consist of many interdependent objects (we shall only deal with linear systems). Then, there may be multiple pure frequencies, known as its normal frequencies, that the system can oscillate at.

10.4.1 *Decoupling*

Consider the common example below. Two equal masses m are connected by three springs with equal spring constants k to each other and to the walls adjacent to them. Solve for the displacements of the masses as a function of time and the normal frequencies of oscillation if the masses were given a slight initial displacement.

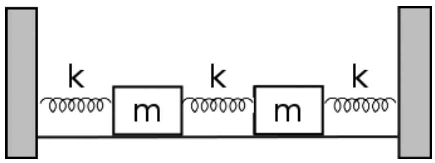


Figure 10.10: Coupled masses

Let x_1 and x_2 be the displacements of the two masses from their respective equilibrium positions, with the rightwards direction taken to be positive. Writing the equation of motion for each individual mass,

$$m\ddot{x}_1 = -kx_1 + k(x_2 - x_1),$$
$$m\ddot{x}_2 = -kx_2 - k(x_2 - x_1).$$

Be careful with the signs here. A reliable way to determine the signs would be to envision the physical scenario. Supposing that $x_2 > x_1$, this would physically mean that the middle spring has been stretched. Thus, the middle spring will pull the first mass towards the right ($+k(x_2 - x_1)$ in the first equation) and pull the second mass towards the left ($-k(x_2 - x_1)$ in the second equation).

Observing the equations of motion, we realise that they are "coupled" in the sense that the way in which the state of one object evolves depends on the state of the other object. To solve this pair of equations, we have to decouple them. Adding the two equations,

$$(\ddot{x}_1 + \ddot{x}_2) = -\frac{k}{m}(x_1 + x_2)$$
$$\implies x_1 + x_2 = A\sin\left(\sqrt{\frac{k}{m}}t + \phi_1\right).$$

Subtracting the second equation from the first,

$$(\ddot{x}_1 - \ddot{x}_2) = -\frac{3k}{m}(x_1 - x_2)$$
$$\implies x_1 - x_2 = B\sin\left(\sqrt{\frac{3k}{m}}t + \phi_2\right).$$

Thus,

$$x_1 = \frac{A}{2}\sin\left(\sqrt{\frac{k}{m}}t + \phi_1\right) + \frac{B}{2}\sin\left(\sqrt{\frac{3k}{m}}t + \phi_2\right),$$

$$x_2 = \frac{A}{2}\sin\left(\sqrt{\frac{k}{m}}t + \phi_1\right) - \frac{B}{2}\sin\left(\sqrt{\frac{3k}{m}}t + \phi_2\right).$$

We see that the two normal frequencies are $\sqrt{\frac{k}{m}}$ and $\sqrt{\frac{3k}{m}}$. The normal modes represent the possible forms of pure-frequency motions and are represented in terms of vectors. The above solution can be represented in matrices as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \frac{A}{2}\sin\left(\sqrt{\frac{k}{m}}t + \phi_1\right) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \frac{B}{2}\sin\left(\sqrt{\frac{3k}{m}}t + \phi_2\right).$$

Then, $(1, 1)$ and $(1, -1)$ are the normal modes of this motion. Let us examine the physical meaning of these modes. With regard to the first normal mode $(1, 1)$, the displacements of the two masses from their equilibrium positions are identical. Then, the middle spring is not stretched or compressed during the entire motion and can be effectively removed — leading to the normal frequency $\sqrt{\frac{k}{m}}$. Next, the normal mode $(1, -1)$ corresponds to displacements of equal magnitude and opposite direction. Then, the middle spring is stretched or compressed twice as much as the displacements of the masses. This, in combination with the springs attached to the walls, causes each mass to be effectively attached to a spring of spring constant $3k$ — implying that its corresponding normal frequency is $\sqrt{\frac{3k}{m}}$.

10.4.2 *General Solution*

In the general case of linear, second-order, coupled simple-harmonic differential equations, the specific way of multiplying equations by constants and adding them in order to successfully decouple them is difficult to spot. Therefore, a general solution would be ideal. With n variables $(x_1, x_2, \ldots, x_n)$, we generally have the following set of equations:

$$\ddot{x}_1 = c_{11}x_1 + c_{12}x_2 + \cdots + c_{1n}x_n$$

$$\ddot{x}_2 = c_{21}x_1 + c_{22}x_2 + \cdots + c_{2n}x_n$$

$$\vdots$$

$$\ddot{x}_n = c_{n1}x_1 + c_{n2}x_2 + \cdots + c_{nn}x_n,$$

where c_{ij} is the constant in the ith row and jth column. We let $\boldsymbol{X}$ be the $n \times 1$ matrix

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

and define A as the $n \times n$ matrix

$$\boldsymbol{A} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & & & \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix}.$$

We can then represent the set of equations compactly by

$$\ddot{\boldsymbol{X}} = \boldsymbol{A}\boldsymbol{X}.$$

Since they are coupled equations, we can try to guess that the solutions for the various x_i's all have the same angular frequency. As long as we can determine a general solution with $2n$ constants to accommodate the $2n$ initial conditions, we are done. Concretely, we guess

$$\boldsymbol{X} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} e^{i\omega t} = \boldsymbol{u}e^{i\omega t}$$

$$\implies \ddot{\boldsymbol{X}} = -\boldsymbol{u}\omega^2 e^{i\omega t}.$$

Substituting these back into the matrix equation, we get

$$-\boldsymbol{u}\omega^2 e^{i\omega t} = \boldsymbol{A}\boldsymbol{u}e^{i\omega t}$$

$$\boldsymbol{A}\boldsymbol{u} = -\omega^2 \boldsymbol{u}. \tag{10.13}$$

Again, any linear combination of the solutions obtained by guessing is also a solution; the general solution is a linear combination of all the linearly independent solutions. Now, notice that if $\boldsymbol{u}e^{i\omega t}$ is a solution, $\boldsymbol{u}e^{-i\omega t}$ is also a solution (as they both result in the same Eq. (10.13). Therefore, instead of including the complex conjugates such as $\boldsymbol{u}e^{-i\omega t}$ in the general solution, we can simply write the general solution as the real component of the linear combination of solutions with positive, imaginary exponents (i.e. $\boldsymbol{u}e^{i\omega t}$) as $\boldsymbol{X}$ must be real at all times. At this juncture, we proceed with a short linear algebra interlude.

Eigenvectors

Let $\boldsymbol{A}$ be a $n \times n$ matrix. A non-null vector $\boldsymbol{u}$ is known as an eigenvector of $\boldsymbol{A}$ if

$$\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}$$

for some scalar λ. λ is termed an eigenvalue of $\boldsymbol{A}$ and $\boldsymbol{u}$ is known as an eigenvector associated with the eigenvalue λ. The eigenvalues can be determined as follows.

$$\boldsymbol{A}\boldsymbol{u} - \lambda \boldsymbol{u} = \mathbf{0}$$

$$(\boldsymbol{A} - \lambda \boldsymbol{I})\,\boldsymbol{u} = \mathbf{0},$$

where $\mathbf{0}$ is a vector of zeroes with n rows and $\boldsymbol{I}$ is the identity matrix of order n. The identity matrix of order n is a $n \times n$ square[4] matrix whose top-left to bottom-right diagonal entries are one — all other entries are zero. As its nomenclature implies, multiplying a square matrix $\boldsymbol{X}$ by the identity matrix of the same dimensions simply returns $\boldsymbol{X}$ ($\boldsymbol{X}\boldsymbol{I} = \boldsymbol{X}$ and $\boldsymbol{I}\boldsymbol{X} = \boldsymbol{X}$). Therefore, we have simply expressed $\boldsymbol{u} = \boldsymbol{I}\boldsymbol{u}$ in writing the second equation.

Now, consider the following definition: an inverse $\boldsymbol{X}^{-1}$ of a square matrix $\boldsymbol{X}$ is defined as a matrix such that the matrix multiplications $\boldsymbol{X}\boldsymbol{X}^{-1} = \boldsymbol{I}$ and $\boldsymbol{X}^{-1}\boldsymbol{X} = \boldsymbol{I}$ (i.e. they yield the identity matrix of the same dimensions as $\boldsymbol{X}$). Suppose that an inverse of $(\boldsymbol{A} - \lambda \boldsymbol{I})$ exists. Then by multiplying this inverse to both sides of the previous equation, we obtain the trivial solution

$$\boldsymbol{u} = \mathbf{0},$$

which is contrary to what we want, as an eigenvector is not a null vector by definition, and the null case is not physically meaningful in the case of coupled oscillations. Thus, in order for non-trivial solutions of $\boldsymbol{u}$ to exist, $(\boldsymbol{A} - \lambda \boldsymbol{I})$ must be non-invertible or singular. In linear algebra, this is equivalent to saying that the determinant of this term is zero.

[4]A square matrix is simply one with an identical number of rows and columns.

Determinants

The determinant of a $n \times n$ square matrix, $\boldsymbol{X}$, is a quantity that can be computed recursively as follows. Given that

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix},$$

$$\det(\boldsymbol{X}) = \begin{cases} x_{11} & \text{if} \quad n = 1 \\ x_{i1}Y_{i1} + x_{i2}Y_{i2} + \cdots + x_{in}Y_{in} & \text{for} \quad n \geq 2, \end{cases}$$

where i refers to that particular row. Y_{ij} is defined as

$$Y_{ij} = (-1)^{i+j}\det(\boldsymbol{Z}_{ij}).$$

$\boldsymbol{Z}_{ij}$ is the matrix obtained by removing the ith row and jth column from $\boldsymbol{X}$, i.e.

$$\boldsymbol{Z}_{ij} = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1(j-1)} & x_{1(j+1)} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2(j-1)} & x_{2(j+1)} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(i-1)1} & x_{(i-1)2} & \dots & x_{(i-1)(j-1)} & x_{(i-1)(j+1)} & \dots & x_{(i-1)n} \\ x_{(i+1)1} & x_{(i+1)2} & \dots & x_{(i+1)(j-1)} & x_{(i+1)(j+1)} & \dots & x_{(i+1)n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{n(j-1)} & x_{n(j+1)} & \dots & x_{nn} \end{pmatrix}.$$

The recursive definition of $\det(\boldsymbol{X})$ for $n \geq 2$ above is known as the co-factor expansion along row i, where i is an arbitrary integer $1 \leq i \leq n$. In fact, the determinant can also be calculated via a co-factor expansion along any column j.

$$\det(\boldsymbol{X}) = \sum_{i=1}^{N} x_{ij}Y_{ij} \quad \text{for} \quad n \geq 2.$$

Let us now evaluate the determinants of two concrete examples to clarify this esoteric definition. The most common form of matrices would be the

2×2 matrix. If we let $\boldsymbol{X}$ be

$$\boldsymbol{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then by performing a co-factor expansion along the first row,

$$\det(\boldsymbol{X}) = aY_{11} + bY_{12} = a \cdot (-1)^{1+1}|d| + b(-1)^{1+2}|c| = ad - bc,$$

where the vertical lines denote taking the determinant of the matrix they enclose.

Problem: Determine the determinant of the following matrix

$$\boldsymbol{X} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Performing a co-factor expansion along the first row,

$$\begin{aligned} \det(\boldsymbol{X}) &= 1 \cdot \begin{vmatrix} 3 & 0 \\ 1 & 2 \end{vmatrix} - 2 \cdot \begin{vmatrix} 1 & 0 \\ 0 & 2 \end{vmatrix} \\ &= (3 \cdot 2 - 1 \cdot 0) - 2 \cdot (1 \cdot 2 - 0 \cdot 0) \\ &= 2. \end{aligned}$$

Incidentally, there is an efficient memorization scheme for the determinant of a 3×3 matrix known as Sarrus' rule.

$$\begin{vmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{vmatrix} = \begin{matrix} x_{11}x_{22}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} \\ -x_{31}x_{22}x_{13} - x_{32}x_{23}x_{11} - x_{33}x_{21}x_{12} \end{matrix}$$

This sum can be visualized by replicating the first two columns on the right of the original block of numbers and taking the sum of the products along the bolded diagonals, minus the sum of the products along the dashed diagonals in Fig 10.11.

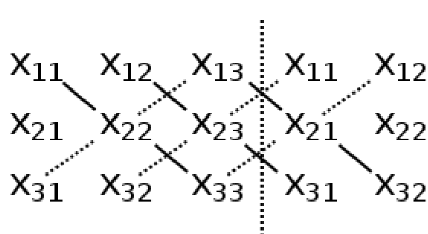


Figure 10.11: Sarrus' rule

Evaluating Eigenvalues and Normal Frequencies

Returning to our main topic of eigenvectors, the eigenvalues associated with a matrix $\boldsymbol{A}$ can be computed by setting the determinant of $\boldsymbol{A} - \lambda \boldsymbol{I}$ to be zero, so

$$\det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right) = 0.$$

This will generate a nth order polynomial for λ which has n roots by the fundamental theorem of algebra. In the case of coupled oscillators, by observing Eq. (10.13), we have

$$\left(\boldsymbol{A} + \omega^2 \boldsymbol{I}\right) \boldsymbol{u} = \boldsymbol{0}, \tag{10.14}$$

which has non-trivial solutions only if

$$\det\left(\boldsymbol{A} + \omega^2 \boldsymbol{I}\right) = 0.$$

That is, the squared negative of the normal frequencies are the eigenvalues of the matrix $\boldsymbol{A}$. Let us consider the specific spring-mass oscillators in the previous section. The equations of motion produced are

$$\ddot{x}_1 = -\frac{2k}{m}x_1 + \frac{k}{m}x_2,$$
$$\ddot{x}_2 = \frac{k}{m}x_1 - \frac{2k}{m}x_2.$$

Therefore, the matrix $\boldsymbol{A}$ is

$$\boldsymbol{A} = \begin{pmatrix} -\frac{2k}{m} & \frac{k}{m} \\ \frac{k}{m} & -\frac{2k}{m} \end{pmatrix}.$$

Then, we require

$$\begin{vmatrix} -\frac{2k}{m} + \omega^2 & \frac{k}{m} \\ \frac{k}{m} & -\frac{2k}{m} + \omega^2 \end{vmatrix} = 0$$
$$\left(-\frac{2k}{m} + \omega^2\right)^2 - \left(\frac{k}{m}\right)^2 = \left(\omega^2 - \frac{3k}{m}\right)\left(\omega^2 - \frac{k}{m}\right) = 0$$
$$\omega^2 = \frac{k}{m} \quad \text{or} \quad \frac{3k}{m}.$$

Thus, the eigenvalues of $\boldsymbol{A}$ are $-\frac{k}{m}$ and $-\frac{3k}{m}$ while the normal frequencies are $\sqrt{\frac{k}{m}}$ and $\sqrt{\frac{3k}{m}}$.

Evaluating Eigenvectors and Normal Modes

Now that we have computed the eigenvalues of $\boldsymbol{A}$, we can now determine the eigenvectors associated with an eigenvalue λ_i by substituting $\lambda = \lambda_i$ back into the equation $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{u} = \boldsymbol{0}$. Then, we can solve the resultant matrix equation (some variables will still be expressed in terms of the others) to obtain a general solution for $\boldsymbol{u}$ that satisfies the equation. This general solution, which is expressed in terms of a linear combination of independent vectors, is known as the eigenspace associated with the eigenvalue λ_i. The eigenspace is usually denoted as $\boldsymbol{E}_{\lambda_i}$ but we shall denote it as $\boldsymbol{E}_{-\lambda_i}$ (as $\lambda_i = -\omega_i^2$ where ω_i is the ith normal frequency) for our purposes. The independent vectors which appear in the linear combination are the basis eigenvectors associated with the eigenvalue λ_i, as substituting any linear combination of them for $\boldsymbol{u}$ in $\boldsymbol{A}\boldsymbol{u}$ would result in $\lambda_i \boldsymbol{u}$. Furthermore, in the context of coupled oscillators, the basis eigenvectors associated with eigenvalue $\lambda_i = -\omega_i^2$ turn out to be the normal modes associated with the normal frequency ω_i. Do not worry too much about what these terms mean for now and consider the following specific example. In the case of the coupled spring-mass oscillators, we substitute the various values for ω^2 that we have found, into Eq. (10.14).

When $\lambda_1 = -\omega_1^2 = -\frac{k}{m}$, we obtain

$$\begin{pmatrix} -\frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & -\frac{k}{m} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0.$$

Solving gives

$$u_1 = u_2.$$

Therefore, the eigenspace for $\boldsymbol{u}$ associated with λ_1 is the collection of vectors

$$\boldsymbol{E}_{\frac{k}{m}} = C_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

where C_1 is a scalar. That is, any vector $\boldsymbol{u}$ of this form would be an eigenvector associated with the eigenvalue λ_1. Evidently, the only basis eigenvector associated with the eigenvalue λ_1 is $(1, 1)$. Similarly, when $\lambda_2 = -\omega_2^2 = -\frac{3k}{m}$,

$$\begin{pmatrix} \frac{k}{m} & \frac{k}{m} \\ \frac{k}{m} & \frac{k}{m} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 0$$

$$u_2 = -u_1.$$

The eigenspace associated with λ_2 is the collection of vectors

$$\boldsymbol{E}_{\frac{3k}{m}} = C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

for some scalar C_2. The only basis eigenvector associated with the eigenvalue λ_2 is $(1, -1)$. Finally, the general solution for the displacements of the masses, $\boldsymbol{X}$, is obtained by concatenating the various $\boldsymbol{E}_{\omega_i^2} e^{i\omega_i t}$'s. We will not include $\boldsymbol{E}_{\omega_i^2} e^{-i\omega_i t}$ as we will take the real component of the combination later to obtain the physical solution for $\boldsymbol{X}$ (see paragraph below Eq. (10.13)). The expression obtained from patching is

$$\boldsymbol{E}_{\frac{k}{m}} e^{i\sqrt{\frac{k}{m}}t} + \boldsymbol{E}_{\frac{3k}{m}} e^{i\sqrt{\frac{3k}{m}}t} = C_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{i\sqrt{\frac{k}{m}}t} + C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{i\sqrt{\frac{3k}{m}}t}.$$

If we let $C_1 = D_1 e^{i\phi_1}$ and $C_2 = D_2 e^{i\phi_2}$ where D_1, D_2, ϕ_1 and ϕ_2 are real constants, taking the real component of the above expression yields

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} D_1 \cos\left(\sqrt{\frac{k}{m}}t + \phi_1\right) + \begin{pmatrix} 1 \\ -1 \end{pmatrix} D_2 \cos\left(\sqrt{\frac{3k}{m}}t + \phi_2\right).$$

As seen from above, the basis eigenvectors associated with eigenvalue λ_i now function as the normal modes of the normal frequency ω_i. The above expression is the most general solution for $\boldsymbol{X}$ as we have 4 constants to accommodate the 4 initial conditions (positions and velocities of both masses). As a last remark, this method of finding the eigenvalues is not foolproof. If there are repeated eigenvalues, we may not be able to find sufficient linearly independent solutions. Then, we would need to guess other forms of solutions. In the specific case where $\omega^2 = 0$ is a possibility, we should guess polynomials of degree one (i.e. $\boldsymbol{X} = \boldsymbol{u}(c_0 + c_1 t)$) as $\omega^2 = 0$ insinuates that the second derivative of $\boldsymbol{X}$ is zero.

Problems

Simple Harmonic Motion

1. *Pendulum Clock**

A pendulum clock, which has a period of one second when connected to a fixed pivot, is attached to the ceiling of a lift at rest. The lift then undergoes an upwards acceleration a for t seconds. Immediately afterwards, it is slowed down with deceleration a until it stops. Would the clock still be accurate at this juncture? For instance, if the time taken for the whole journey is 10s but the pendulum only oscillates 9 times, the clock would be slower by 1s and is no longer accurate.

2. *Dropping a Mass**

Consider a spring-mass system of mass m and spring constant k on a frictionless, horizontal table. If the initial amplitude is A and another mass m is dropped vertically onto the oscillating mass and sticks with it when its displacement is $\frac{A}{2}$, determine the final amplitude of oscillation A'.

3. *Kinematic Quantities**

Given that the speeds of an oscillating particle at displacements x_1 and x_2 are v_1 and v_2 respectively, determine the amplitude and angular frequency of the oscillation.

4. *Physical Pendulum**

A Physics student measures the period of an arbitrary physical pendulum about a certain pivot to be T. Then, he identifies another pivot on the opposite side of the center of mass that gives the same period. If the two points are separated by a distance l, can he determine the gravitational field strength g of the Earth, assuming that it is uniform throughout the pendulum?

5. *Cavendish Experiment**

The Cavendish experiment was performed to determine the universal gravitational constant G. Two identical small balls of mass m are connected by a light rod with length L and lie on a frictionless table. The center of the rod is connected to the ceiling via a vertical torsion wire. The torsion constant of the wire, κ, is defined as the restoring torque per unit angular twist of the wire.

(1) Find the period T of this torsion pendulum in terms of the above parameters (besides G) when the rod is rotated.
(2) Now, place two identical large balls of mass M at two diametrically opposite points on the perimeter of a circle of diameter L about the center of the rod (i.e. the small balls lie on the same circle). When the system is at equilibrium, the rod has rotated an angle θ and the distance between the center of a small ball and its adjacent large ball is $r \ll L$. Determine G in terms of L, r, M, T and θ.
(3) Suppose that the small balls are perturbed by a small angle from the equilibrium position. Will they oscillate about the equilibrium position? If so, determine the angular frequency of such oscillations in terms of κ, θ, m, L and r.

6. *Particle in Potential**

A particle of mass m is acted on by a one-dimensional potential energy given by

$$U(x) = U_0(-ax^2 + bx^4),$$

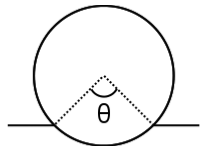
where U_0, a and b are positive constants. Determine the equilibrium x-coordinates of the particle and classify them as stable or unstable. If an equilibrium position is stable, determine the angular frequency of small oscillations about it.

7. *V-Shape Rails**

Two particles of common mass m are constrained to move along two rails, which subtend an angle 2θ, that form a V-shape. The particles are connected by a spring with spring constant k. What is the angular frequency of oscillations for the motion where the spring remains perpendicular to the symmetrical axis of the rails?

8. *Floating Cylinder**

A cylinder of density d, radius r and length l is floating on water of density ρ as shown in the diagram on the next page. Write an expression for the equilibrium value of the angle θ, subtended by the wetted portion of the cylinder, as labeled in the diagram on the next page. If the cylinder is now pressed slightly downwards, determine the angular frequency of small oscillations. Feel free to express your answer in terms of the equilibrium angle θ.

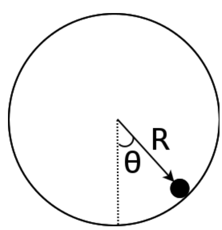


9. *Two Circles**

A small circle of radius r is attached to the circumference of a large circle of radius R. If the surface mass density of the circles is σ, determine the angular frequency of small oscillations about the equilibrium position if the center of the large circle is pivoted.

10. *Two Spheres***

A spherical ball of radius r, mass m and uniform mass density rolls without slipping in the interior of a sphere with mass M, radius R and uniform mass density near the bottom of the sphere, solely in the θ direction. The large sphere cannot translate but it may rotate. What is the angular frequency of small oscillations of the ball about the bottom of the large sphere?

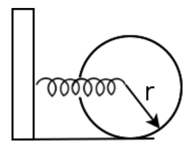


11. *Masses and String***

A mass m is undergoing circular motion about a hole on a horizontal table at radius r_0. A string, passing through the hole, is attached to m and another mass M which hangs vertically. If mass m is given a slight radial push, determine the angular frequency of small oscillations in the radial direction.

12. *Non-slip Oscillation***

Referring to the figure on the next page, a cylinder of mass m and radius r lies with its cylindrical axis in the plane of the horizontal ground. A spring of spring constant k and relaxed length l is attached to the center of the cylinder at one end and a fixed wall at the other end. If the cylinder does not slip with the ground, determine the angular frequency of oscillations.



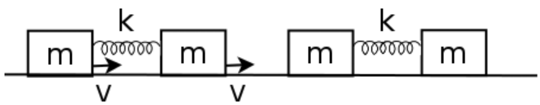
13. *Particle in a Sphere****

A particle of mass m is currently undergoing circular motion at angular frequency $\omega_0 > \sqrt{\frac{g}{R}}$ in the interior of a massive sphere of radius R. Let θ be the angle between the vertical axis, passing through the center of the sphere and pointing downwards, and the position of the particle in spherical coordinates, taking a positive value in the anti-clockwise direction. Suppose that the particle is given a slight push in the θ direction, determine the angular frequency of small oscillations in the θ direction.

Damped and Coupled Oscillators

14. *Colliding Couples***

Two point masses of mass m are connected by a spring of spring constant k and relaxed length l. The two masses both have an initial velocity v and the spring between them stays at its relaxed length. These masses then travel towards an identical set-up (consisting of two masses connected by a spring) on a frictionless, horizontal table. If these four masses are aligned and undergo perfectly elastic, head-on collisions, determine the equations of motion of the masses after the first collision and before the second collision. Determine the elapsed time between the first and second collisions and show that there will only be a total of two collisions.



15. *Colliding Masses***

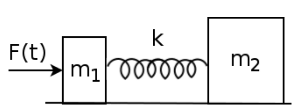
A particle of mass M approaches two initially stationary particles of common mass $m = 2$kg that are connected by a spring of spring constant $k = 1$N/m, at an initial velocity v_0. The collision is one-dimensional, elastic and instantaneous. Determine the minimum value of M for which M will again collide with the system comprising the other particles. As you will obtain a transcendental equation, an approximate value is fine. How much time will elapse between the two collisions for this particular value of M?

16. *Spring-Mass with Friction***

A horizontal spring-mass system is placed on a rough table, with one end of the spring attached to a vertical wall. The massless spring has spring constant k and rest length l_0 while the load has mass m and static and kinetic friction coefficients μ relative to the table. Initially, the mass m is placed right next to the wall such that the length of the spring is virtually zero. Subsequently, m is released such that it begins to undergo a one-dimensional oscillatory motion. If we define the origin to be at the fixed end of the spring and the x-axis to be along the direction of motion of m, sketch the graph of the x-coordinate x of the mass against the elapsed time t. Thus, determine the total number of oscillation cycles that m completes before coming to a stop.

17. *Stabilizer***

A heavy bob is often used to stabilize buildings in the event of earthquakes. Let us consider a related problem. Two masses m_1 and m_2 are stationary on a horizontal, frictionless plane and are connected by a spring of spring constant k. Suppose a force $F(t) = f\cos\omega t$ is exerted on m_1, in the direction of the line joining the two masses. Determine the value of k for which the particular solution to the equation of motion of m_1 yields an oscillation of zero amplitude. There is no damping.



18. *Masses on Hoop****

Three masses, one with mass m, and two with mass $2m$ are constrained to move along a massive circular hoop of radius R. Three springs that are wrapped around the hoop connect adjacent masses. The springs between mass m and the two masses $2m$ have spring constant $2k$ while the spring between the two masses of mass $2m$ has spring constant k. Find the normal modes of oscillation and the displacements of the masses from their equilibrium positions as functions of time under arbitrary initial conditions.

Solutions

1. Pendulum Clock*

The frequency of a pendulum, under an effective gravity g_{eff}, is $f = \frac{1}{2\pi}\sqrt{\frac{g_{eff}}{l}}$. When the lift is accelerating upwards, the pendulum experiences an inertial force ma downwards where m is its mass and hence lives in a world with effective gravity $g + a$. Similarly, when the lift is decelerating, the effective gravity is $g - a$. Since $f \propto \sqrt{g_{eff}}$, the elapsed time that the pendulum clock would have recorded during this experiment is

$$t \cdot \sqrt{\frac{g+a}{g}} + t \cdot \sqrt{\frac{g-a}{g}} \neq 2t,$$

where $2t$ is the actual time elapsed. Therefore, the clock is no longer accurate.

2. Dropping a Mass*

The total mechanical energy is initially

$$E = \frac{1}{2}kA^2.$$

When the mass m is at a state with displacement $\frac{A}{2}$, its potential and kinetic energies are respectively

$$U = \frac{1}{8}kA^2,$$
$$T = E - U = \frac{3}{8}kA^2.$$

Thus, its speed at this instant is

$$v = \sqrt{\frac{3kA^2}{4m}}.$$

The final speed of the two masses after the collision is given by the conservation of momentum to be

$$v' = \sqrt{\frac{3kA^2}{16m}}.$$

Thus, the final kinetic energy is

$$T' = \frac{1}{2} \cdot 2mv'^2 = \frac{3kA^2}{16}.$$

The total mechanical energy afterwards is

$$E = U + T' = \frac{5kA^2}{16}.$$

At the amplitude A', the kinetic energy of the mass is zero. Thus,

$$\frac{1}{2}kA'^2 = E$$

$$A' = \sqrt{\frac{5}{8}}A.$$

3. Kinematic Quantities*

$$|v_1| = \omega\sqrt{A^2 - x_1^2},$$

$$|v_2| = \omega\sqrt{A^2 - x_2^2}.$$

Dividing the first equation by the second and squaring, we get

$$\frac{v_1^2}{v_2^2} = \frac{A^2 - x_1^2}{A^2 - x_2^2}.$$

Solving,

$$A = \sqrt{\frac{v_2^2 x_1^2 - v_1^2 x_2^2}{v_2^2 - v_1^2}}.$$

Substituting this expression for A into either of the first two equations,

$$\omega = \sqrt{\frac{v_2^2 - v_1^2}{x_1^2 - x_2^2}}.$$

4. Physical Pendulum*

Applying Eq. (10.6), the period of a physical pendulum is

$$T = 2\pi\sqrt{\frac{I_{pivot}}{Mr_{CM}g}},$$

where I_{pivot} is the moment of inertia of the pendulum about the pivot, M is the total mass of the physical pendulum and r_{CM} is the distance between the pivot and the center of mass of the pendulum. Let the distances between the two pivots and the center of mass be r and $l-r$ respectively. Then,

$$2\pi\sqrt{\frac{I_{CM}+Mr^2}{Mrg}}=T,$$

$$2\pi\sqrt{\frac{I_{CM}+M(l-r)^2}{M(l-r)g}}=T,$$

where I_{CM} is the moment of inertia of the pendulum about its center of mass. Eliminating I_{CM},

$$\frac{T^2Mrg}{4\pi^2}-Mr^2=\frac{T^2M(l-r)g}{4\pi^2}-M(l-r)^2$$

$$\implies g=\frac{4\pi^2 l}{T^2}.$$

5. Cavendish Experiment*

Due to symmetry, the center of the rod is the instantaneous center of rotation. The moment of inertia of the rod and the small balls about the center is $I=2\cdot m\frac{L^2}{4}=\frac{mL^2}{2}$. When the rod has rotated for an angle θ, the restoring torque is $-\kappa\theta$. Therefore, the equation of motion of the rod is

$$I\ddot{\theta}=-\kappa\theta$$

$$\ddot{\theta}=-\frac{2\kappa}{mL^2}\theta,$$

which indicates a simple harmonic motion of period

$$T=2\pi\sqrt{\frac{mL^2}{2\kappa}}.$$

For the second part, the torque produced by the gravitational force of the large balls balances the torsion torque at equilibrium. Note that we only consider the gravitational force on a small ball due to the nearer large ball. The other gravitational force is comparatively negligible, since $r\ll L$. Balancing torques,

$$\frac{GMm}{r^2}\cdot L=\kappa\theta$$

$$G=\frac{\kappa\theta r^2}{MmL}=\frac{2\pi^2r^2L\theta}{MT^2}.$$

Let $\theta_M=\theta_0+\frac{2r}{L}$ denote the angular position of the large balls, relative to the original position of the rod. θ_0 is the angle θ in the previous section (we reserve the variable θ for the following definition). When the rod has experienced an angular displacement θ ($\theta_M - \theta \ll \theta_M$), the separation between the centers of neighboring small and large balls is approximately $\frac{L}{2}(\theta_M - \theta)$. Therefore, the equation of motion of the rod is

$$I\ddot{\theta} = 2 \cdot \frac{GMm}{\frac{L^2}{4}(\theta_M - \theta)^2} \cdot \frac{L}{2} - \kappa\theta$$

$$\ddot{\theta} = \frac{8GM}{L^3(\theta_M - \theta)^2} - \frac{2\kappa}{mL^2}\theta.$$

When $\theta = \theta_0$, the system is in equilibrium such that

$$\frac{8GM}{L^3 \cdot \frac{4r^2}{L^2}} - \frac{2\kappa}{mL^2}\theta_0 = 0.$$

Now, suppose $\theta = \theta_0 + \varepsilon$. The equation of motion becomes

$$\ddot{\varepsilon} = \frac{8GM}{L^3\left(\frac{2r}{L} - \varepsilon\right)^2} - \frac{2\kappa}{mL^2}(\theta_0 + \varepsilon)$$

$$\ddot{\varepsilon} = \frac{2GM}{r^2L\left(1 - \frac{L\varepsilon}{2r}\right)^2} - \frac{2\kappa}{mL^2}(\theta_0 + \varepsilon).$$

Performing a first-order binomial expansion,

$$\begin{aligned}\ddot{\varepsilon} &= \frac{2GM}{r^2L}\left(1 + \frac{L\varepsilon}{r}\right) - \frac{2\kappa}{mL^2}(\theta_0 + \varepsilon) \\ &= -\left(\frac{2\kappa}{mL^2} - \frac{2GM}{r^3}\right)\varepsilon = -\left(\frac{2\kappa}{mL^2} - \frac{2\kappa\theta_0}{mLr}\right)\varepsilon,\end{aligned}$$

where we have performed the cancellation of some terms based on the previous equilibrium equation and substituted $\frac{GM}{r^2} = \frac{\kappa\theta_0}{mL}$. The above indicates a simple harmonic motion of angular frequency

$$\omega = \sqrt{\frac{2\kappa(r - \theta_0 L)}{mL^2 r}}$$

if $r > \theta_0 L$. Otherwise, the rod will not exhibit simple harmonic motion.

6. Particle in Potential*

Since a conservative force is the negative potential energy gradient, the equilibrium positions correspond to the locations where $U'(x) = 0$.

$$U'(x) = U_0(-2ax + 4bx^3) = U_0 x(4bx^2 - 2a).$$

The equilibrium positions are thus $x = 0$, $x = -\sqrt{\frac{a}{2b}}$ and $x = \sqrt{\frac{a}{2b}}$. Computing the second derivative,

$$U''(x) = U_0(12bx^2 - 2a).$$

Since $U''(0) = -2aU_0 < 0$, the equilibrium position $x = 0$ corresponds to a potential energy maximum which indicates that a slight deviation tends to be amplified by the conservative force (which is directed towards lower values of potential energy). Hence, $x = 0$ is unstable but on the other hand, $U''(\pm\sqrt{\frac{a}{2b}}) = 4aU_0 > 0$ which indicates that $x = -\sqrt{\frac{a}{2b}}$ and $x = \sqrt{\frac{a}{2b}}$ are stable equilibria. The angular frequency of oscillations about these positions is

$$\omega = \sqrt{\frac{U''\left(\pm\sqrt{\frac{a}{2b}}\right)}{m}} = \sqrt{\frac{4aU_0}{m}}.$$

7. V-Shape Rails*

Suppose that both masses are shifted along the rails by a displacement x from their equilibrium positions. The spring would have stretched or contracted by an additional $2x\sin\theta$, beyond its length when the two masses are at equilibrium. Therefore, the equation of motion of one mass at this juncture is

$$m\ddot{x} = -2k\sin\theta x \cdot \sin\theta,$$

where we multiply by $\sin\theta$ to obtain the component of force along the rail that it lies along. Since

$$\ddot{x} = -\frac{2k\sin^2\theta}{m}x,$$

the angular frequency of small oscillations is

$$\omega = \sqrt{\frac{2k\sin^2\theta}{m}}.$$

An alternative method for this question would start with the potential energy of the system when each mass is at a distance x from the point of connection of the rails.

$$U(x) = \frac{1}{2}k(2x\sin\theta - l_0)^2,$$

where l_0 is the rest length of the spring. The first derivative is

$$U'(x) = k(2x\sin\theta - l_0)\cdot 2\sin\theta,$$

which shows that the equilibrium x-coordinate is $\frac{l_0}{2\sin\theta}$. The second derivative of this is

$$U''(x) = 4k\sin^2\theta.$$

Therefore, the angular frequency of small oscillations about the equilibrium position is

$$\omega = \sqrt{\frac{U''\left(\frac{l_0}{2\sin\theta}\right)}{2m}} = \sqrt{\frac{2k\sin^2\theta}{m}}.$$

Note that we have to use $2m$ instead of m here as $U(x)$ is the potential energy of the entire system.

8. Floating Cylinder*

The volume of the cylinder submerged in water is the area of the sector (multiplied by l) minus the area of the isosceles triangle, with sides r that subtend angle θ (multiplied by l).

$$V = \frac{\theta}{2}r^2 l - \frac{1}{2}r^2 l\sin\theta.$$

The cylinder is in equilibrium when the upthrust balances its weight.

$$\rho\left(\frac{\theta}{2}r^2 l - \frac{1}{2}r^2 l\sin\theta\right)g = \pi r^2 l d g$$

$$\theta - \sin\theta = \frac{2\pi d}{\rho}.$$

When the cylinder is displaced by a vertical small distance ε from its equilibrium position, the net force that it experiences (which opposes its deviation) is equal to ρg multiplied by the change in the submerged volume of the cylinder. The latter is equal to the length of the horizontal chord on the cross-section of the cylinder along the water level, $2r\sin\frac{\theta}{2}$, multiplied by ε

(the vertical displacement) and l. Therefore, the equation of motion of the cylinder is

$$\pi r^2 l d\ddot{\varepsilon} = -2r\sin\frac{\theta}{2}\cdot l\cdot\rho g\varepsilon$$

$$\ddot{\varepsilon} = -\frac{2\rho g\sin\frac{\theta}{2}}{\pi r d}\varepsilon,$$

which indicates a simple harmonic motion of angular frequency

$$\omega = \sqrt{\frac{2\rho g\sin\frac{\theta}{2}}{\pi r d}}.$$

9. Two Circles*

The moment of inertia of the smaller circle about the center of the larger circle is $\frac{1}{2}\sigma\pi r^4+\sigma\pi r^2R^2$ by the parallel axis theorem. Thus, the total moment of inertia of the system about the pivot is

$$I = \frac{1}{2}\sigma\pi R^4 + \frac{1}{2}\sigma\pi r^4 + \sigma\pi r^2R^2 = \frac{1}{2}\sigma\pi(r^2+R^2)^2.$$

The net external torque on this system is that due to the weight of the smaller circle.

$$\tau = -\sigma\pi r^2 gR\sin\theta$$

$$\implies I\ddot{\theta} = -\sigma\pi r^2 gR\sin\theta.$$

Using the small angle approximation $\sin\theta\approx\theta$,

$$\ddot{\theta} = -\frac{2gr^2R}{(r^2+R^2)^2}\theta.$$

The angular frequency of small oscillations about $\theta = 0$ is

$$\omega = \frac{\sqrt{2gr^2R}}{r^2+R^2}.$$

10. Two Spheres**

Let θ be the angle that the line joining the center of the spheres makes with the vertical, and let it be positive in the anti-clockwise direction. Let ϕ and ψ be the angles that the ball and large sphere have rotated about their centers respectively, also positive anti-clockwise. Since $(R-r)\dot{\theta}$ is the velocity of the center of the ball, $(R-r)\dot{\theta}+r\dot{\phi}$ is the velocity of the point on the ball that is in contact with the large sphere. Therefore, the non-slip condition is

$$(R-r)\dot{\theta}+r\dot{\phi}=R\dot{\psi}$$
$$\implies (R-r)\ddot{\theta}+r\ddot{\phi}=R\ddot{\psi}.$$

Let f be the friction force on the ball in the anti-clockwise direction due to the large sphere. Applying Newton's second law to the ball,

$$f-mg\sin\theta=m(R-r)\ddot{\theta}.$$

Applying $\tau=I\alpha$ to the spheres about their respective centers,

$$f=\frac{2}{5}mr\ddot{\phi},$$
$$-f=\frac{2}{5}MR\ddot{\psi}.$$

Then,

$$r\ddot{\phi}=-\frac{M}{m}R\ddot{\psi}.$$

Substituting this into the non-slip condition,

$$R\ddot{\psi}=\frac{m(R-r)}{m+M}\ddot{\theta},$$
$$f=-\frac{2mM(R-r)}{5(m+M)}\ddot{\theta}.$$

Substituting this into the equation obtained from Newton's second law,

$$\frac{m(R-r)(5m+7M)}{5(m+M)}\ddot{\theta}=-mg\sin\theta.$$

Using the small angle approximation $\sin\theta\approx\theta$,

$$\ddot{\theta}=-\frac{5(m+M)g}{(R-r)(5m+7M)}\theta.$$

The angular frequency of small oscillations about $\theta=0$ is thus

$$\omega=\sqrt{\frac{5(m+M)g}{(R-r)(5m+7M)}}.$$

11. Masses and String**

Let T be the tension in the string and let r be the radial coordinate of m with respect to the hole. Then, the equations of motion of m and M are

$$-T = m(\ddot{r} - r\dot{\theta}^2),$$

$$T - Mg = M\ddot{r},$$

by the conservation of string.

$$\implies -Mg = (m+M)\ddot{r} - mr\dot{\theta}^2.$$

Observe that the angular momentum of mass m about the hole is conserved as it only experiences a radial force. Then,

$$L = mr^2\dot{\theta}$$

for some constant L.

$$-Mg = (m+M)\ddot{r} - \frac{L^2}{mr^3}.$$

When $r = r_0$, $\ddot{r} = 0$, hence

$$Mg = \frac{L^2}{mr_0^3}.$$

This will be useful in canceling terms later. Next, express the radial coordinate r as $r_0 + \varepsilon$ where ε is a slight displacement from the equilibrium position. Then,

$$-Mg = (m+M)\ddot{\varepsilon} - \frac{L^2}{mr_0^3\left(1 - \frac{\varepsilon}{r_0}\right)^3}.$$

Performing a binomial expansion,

$$-Mg = (m+M)\ddot{\varepsilon} - \frac{L^2}{mr_0^3}\left(1 - \frac{3\varepsilon}{r_0}\right).$$

Substituting $\frac{L^2}{mr_0^3} = Mg$,

$$\ddot{\varepsilon} = -\frac{3Mg}{(m+M)r_0}\varepsilon.$$

The angular frequency of small oscillations is thus

$$\omega = \sqrt{\frac{3Mg}{(m+M)r_0}}.$$

An alternative method derives the effective potential of the combined system which is

$$U_{eff}(r) = \frac{L^2}{2mr^2} + Mg(r - l),$$

where the first term is associated with the azimuthal motion of m and the second term is the gravitational potential energy of M.

$$U'_{eff}(r) = -\frac{L^2}{mr^3} + Mg,$$

$$U''_{eff}(r) = \frac{3L^2}{mr^4},$$

$$U''_{eff}(r_0) = \frac{3L^2}{mr_0^4} = \frac{3Mg}{r_0}.$$

The angular frequency of small oscillations about $r = r_0$ is then

$$\omega = \sqrt{\frac{U''_{eff}(r_0)}{m + M}} = \sqrt{\frac{3Mg}{(m + M)r_0}}.$$

Be wary that the mass of the combined system is $m + M$ and not m.

12. Non-slip Oscillation**

It is easier to derive the equation of motion by differentiating the conservation of energy equation. Let the x-coordinate of the center of the cylinder be x with respect to the wall. Then, the total mechanical energy of the cylinder is

$$E = \frac{1}{4}mr^2\omega^2 + \frac{1}{2}m\dot{x}^2 + \frac{1}{2}k(x - l)^2.$$

By the non-slip condition, $r\omega = \dot{x}$. Thus,

$$E = \frac{3}{4}m\dot{x}^2 + \frac{1}{2}k(x - l)^2.$$

Differentiating the above by x and using the fact that $\frac{d(\dot{x}^2)}{dx} = 2\ddot{x}$,

$$\ddot{x} = -\frac{2k}{3m}(x - l).$$

Therefore, the angular frequency of oscillations about $x = l$ is

$$\omega = \sqrt{\frac{2k}{3m}}.$$

13. Particle in a Sphere***

We first determine the initial θ coordinate, θ_0, at which the particle can undergo circular motion. Let N be the normal force on the particle by the sphere. The vertical component of the normal force must balance the weight of the particle, so

$$N\cos\theta_0 = mg.$$

The horizontal component must provide the required centripetal force, and

$$\begin{aligned}N\sin\theta_0 &= mR\sin\theta_0\omega_0^2,\\ \cos\theta_0 &= \frac{g}{R\omega_0^2},\\ \sin^2\theta_0 &= \frac{R^2\omega_0^4 - g^2}{R^2\omega_0^4}.\end{aligned}$$

Now, let ϕ be the azimuthal angle of the particle. Then, the total mechanical energy of the particle is

$$E = \frac{1}{2}mR^2\dot{\theta}^2 + \frac{1}{2}mR^2\sin^2\theta\dot{\phi}^2 - mgR\cos\theta.$$

The component of angular momentum of the particle along the vertical direction is conserved.

$$L = mR^2\sin^2\theta\dot{\phi}.$$

Therefore,

$$E = \frac{1}{2}mR^2\dot{\theta}^2 + \frac{L^2}{2mR^2\sin^2\theta} - mgR\cos\theta.$$

This is equivalent to a one-dimensional motion with respect to coordinate $R\theta$ under the effective potential

$$U_{eff} = \frac{L^2}{2mR^2\sin^2\theta} - mgR\cos\theta.$$

Differentiating twice with respect to $R\theta$,

$$U''_{eff} = \frac{d^2U_{eff}}{d(R\theta)^2} = \frac{mg\cos\theta}{R} + \frac{L^2}{mR^4\sin^2\theta} + \frac{3L^2\cos\theta}{mR^4\sin^4\theta},$$

with

$$L = mR^2\sin^2\theta_0\omega_0 = \frac{m(R^2\omega_0^4 - g^2)}{\omega_0^3}.$$

The angular frequency of small oscillations about $\theta = \theta_0$ is

$$\Omega = \sqrt{\frac{U''_{eff}(\theta_0)}{m}} = \sqrt{\frac{R\omega_0^2 + 3g}{R}}.$$

14. Colliding Couples**

We have proven in the chapter on collisions that when two identical masses collide elastically, they simply swap velocities. Let the rightwards direction be the positive x-direction. Label the masses from 1 to 4 rightwards. Then, the velocities of mass 2 and 3 after the first collision are 0 and v respectively. Therefore, we obtain two identical set-ups, displaced by a distance l. We first solve for a system of two identical masses, connected by a spring of spring constant k and relaxed length l, and separated by an initial distance l. The initial velocities of the masses, from left to right, are v and 0 respectively. If we let y_1 and y_2 represent the coordinates of the masses and consider their equations of motion,

$$m\ddot{y}_1 = k(y_2 - y_1 - l),$$
$$m\ddot{y}_2 = -k(y_2 - y_1 - l),$$
$$\ddot{y}_1 - \ddot{y}_2 = \frac{2k}{m}(y_2 - y_1 - l).$$

Substituting $u = y_2 - y_1 - l$,

$$\ddot{u} = -\frac{2k}{m}u,$$
$$y_2 - y_1 = l + A\sin\left(\sqrt{\frac{2k}{m}}t + \phi\right),$$
$$\dot{y}_2 - \dot{y}_1 = A\sqrt{\frac{2k}{m}}\cos\left(\sqrt{\frac{2k}{m}}t + \phi\right).$$

Substituting the initial conditions $y_2 - y_1 = l$ and $\dot{y}_2 - \dot{y}_1 = -v$ when $t = 0$,

$$A\sin\phi = 0,$$
$$A\sqrt{\frac{2k}{m}}\cos\phi = -v.$$

Out of the possible solutions to the above equations, we choose

$$\phi = 0,$$
$$A = -v\sqrt{\frac{m}{2k}},$$

as any set of possible solutions would yield the same result for y_1 and y_2. Then,

$$y_2 - y_1 = l - v\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right).$$

Next, we know from the conservation of momentum (or by adding the two equations of motion together) that

$$\dot{y}_1 + \dot{y}_2 = v$$
$$y_1 + y_2 = vt + c,$$

where c is a constant that depends on the choice of origin. Then,

$$y_1 = \frac{c}{2} + \frac{vt}{2} - \frac{l}{2} + \frac{v}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right),$$

$$y_2 = \frac{c}{2} + \frac{vt}{2} + \frac{l}{2} - \frac{v}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right).$$

Now, apply these results to our system at hand. Let the first collision between the second and third masses occur at $t = 0$ and let the coordinates of the masses be x_1, x_2, x_3 and x_4. Observe that the pairs x_1, x_2 and x_3, x_4 are analogous to y_1, y_2. If the origin is defined at the point of collision, substituting the initial conditions $x_1 = -l$, $x_2 = 0$, $x_3 = 0$ and $x_4 = l$ at $t = 0$ yields

$$x_1 = -l + \frac{vt}{2} + \frac{v}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right),$$

$$x_2 = \frac{vt}{2} - \frac{v}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right),$$

$$x_3 = \frac{vt}{2} + \frac{v}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right),$$

$$x_4 = l + \frac{vt}{2} + \frac{v}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right).$$

For the second and third masses to collide again, $x_2 = x_3$. This implies that the time elapsed between the first and second collisions is

$$t = \pi\sqrt{\frac{m}{2k}}.$$

At this instant (immediately before the second collision), one can easily show that

$$\begin{aligned}\dot{x}_1 &= 0,\\ \dot{x}_2 &= v,\\ \dot{x}_3 &= 0,\\ \dot{x}_4 &= v.\end{aligned}$$

After the second collision, the third and fourth masses will have both acquired velocity v while the first and second masses will be stationary. Effectively, the initial velocities of the first and second masses have been transferred to the third and fourth masses, which then move off. Thus, only two collisions occur.

15. Colliding Masses**

Let A denote the particle of mass m that M collides with. Define the x-axis to be along the direction of motion of this one-dimensional system and the origin at the point of collision. Referring to the results of the previous problem, the x-coordinate of A after the collision is

$$x_A = \frac{ut}{2} + \frac{u}{2}\sqrt{\frac{m}{2k}}\sin\left(\sqrt{\frac{2k}{m}}t\right),$$

where u is its initial velocity directly after the collision. This can be computed as follows. During the collision, the only impulsive force on M and particle A are the normal forces due to each other (the spring does not exert any impulse in the short collision period). Therefore, this is just a one-dimensional collision between two masses M and m. Defining v and u as the final velocities of M and A, we have

$$Mv_0 = Mv + mu.$$

Since the relative velocity reverses during a one-dimensional elastic collision,

$$u - v = v_0.$$

Solving,

$$v = \frac{M-m}{M+m}v_0 = \frac{1-r}{1+r}v_0,$$
$$u = \frac{2M}{M+m} = \frac{2}{1+r}v_0,$$

where $r = \frac{m}{M}$. The x-coordinate of M after the first collision and before the second collision obeys the equation

$$x_M = vt = \frac{1-r}{1+r}v_0 t.$$

After substituting the values of k and m, the condition for M and particle A to collide again is

$$\frac{1-r}{1+r}v_0 t = \frac{v_0}{1+r}t + \frac{v_0}{1+r}\sin t$$
$$\sin t = -rt.$$

To visualize the maximization of r, we can plot the graph of $y(t) = \sin t$. For there to be a solution to the above equation for a particular value of r, the line $y(t) = -rt$ must intersect $y = \sin t$ at least once. The largest value of r (and hence the steepest linear graph) occurs when $y = -rt$ roughly intersects $y = \sin t$ at its first negative peak ($t = \frac{3\pi}{2}$). In this case, the value of r is

$$r = \frac{1}{\frac{3\pi}{2}} = \frac{2}{3\pi} \implies M = 3\pi\text{kg},$$

and the time elapsed between the two collisions is $\frac{3\pi}{2}$s.

16. Spring-Mass with Friction**

At x-coordinate x, the mass experiences the spring force $-k(x-l_0)$ and friction μmg, whose direction depends on the velocity of the mass. In the outbound regime where the velocity of m is positive (i.e. $\dot{x} > 0$), its equation of motion reads

$$m\ddot{x} = -k(x-l_0) - \mu mg$$
$$\implies \ddot{x} = -\frac{k}{m}\left(x - l_0 + \frac{\mu mg}{k}\right).$$

This equation describes a simple harmonic motion about an equilibrium position at $x = l_0 - \frac{\mu mg}{k}$ with angular frequency $\sqrt{\frac{k}{m}}$. To see this more explicitly, introduce a new variable $u = x - l_0 + \frac{\mu mg}{k}$. Then,

$$\ddot{u} = -\frac{k}{m}u,$$

which describes a simple harmonic motion about $u = 0$ — implying that the equilibrium position is $x = l_0 - \frac{\mu mg}{k}$. In the other case where the motion is inbound (i.e. $\dot{x} < 0$), the equation of motion of m is

$$m\ddot{x} = -k(x - l_0) + \mu mg$$

$$\implies \ddot{x} = -\frac{k}{m}\left(x - l_0 - \frac{\mu mg}{k}\right),$$

which indicates a simple harmonic motion about an equilibrium position at $x = l_0 + \frac{\mu mg}{k}$ with angular frequency $\sqrt{\frac{k}{m}}$. Let us now examine the $(n+1)$th oscillation cycle where m begins at zero initial velocity at x-coordinate x_n but tends to gain a positive velocity (we assume that it moves for now). m will first oscillate about the equilibrium position $x = l_0 - \frac{\mu mg}{k}$ until it reaches x-coordinate $x'_n = 2l_0 - \frac{2\mu mg}{k} - x_n$ where it attains zero velocity again (we use a prime to denote the positions of zero velocity that tend to result in a subsequent negative velocity). If m still starts to move at this juncture, it will tend to gain a negative velocity and oscillate about the equilibrium position $x = l_0 + \frac{\mu mg}{k}$ until it reaches x-coordinate $x_{n+1} = 2\left(l_0 + \frac{\mu mg}{k}\right) - x'_n = x_n + \frac{4\mu mg}{k}$ where it possesses zero velocity again — completing an oscillation cycle. m can only possibly stop at the junctures where it attains zero velocity and will do so if the maximum magnitude of static friction exceeds the spring force then. Therefore, the motion of mass m will terminate at the $(n+1)$th cycle (without completing it) if

$$\frac{k|x_n - l_0|}{mg\cos\theta} \leq \mu,$$

or

$$\frac{k|x'_n - l_0|}{mg\cos\theta} \leq \mu.$$

Otherwise, m will complete the $(n+1)$th cycle and proceed with the $(n+2)$th cycle with

$$x_{n+1} = x_n + \frac{4\mu mg}{k},$$

$$x'_{n+1} = x'_n - \frac{4\mu mg}{k}.$$

Plotting the graph of $x(t)$ with $x = 0$ at $t = 0$, we obtain the following (Fig. 10.12) if $kl_0 > \mu mg$ (such that m actually moves initially).

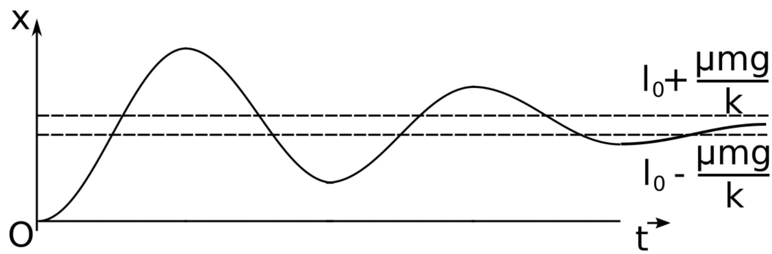


Figure 10.12: Graph of $x(t)$

As the x-coordinate of the particle is increasing (during the ascending portion), it oscillates about the line $x = l_0 - \frac{\mu mg}{k}$. Otherwise when its x-coordinate is decreasing (during the descending portion), it oscillates about $x = l_0 + \frac{\mu mg}{k}$. The t-coordinates of successive peaks and troughs are both separated by the period $T = 2\pi\sqrt{\frac{m}{k}}$. Furthermore, the x-coordinate of a peak is $\frac{4\mu mg}{k}$ lower than its predecessor while the x-coordinate of a trough is $\frac{4\mu mg}{k}$ higher than its predecessor. The motion of m will stop at the first peak or trough that falls into the region between the two horizontal lines. The $(n+1)$th peak and trough occur with x-coordinates

$$x'_n = 2l_0 - \frac{(4n+2)\mu mg}{k},$$
$$x_n = \frac{4n\mu mg}{k}.$$

Observe that x-coordinates of the peaks will always be larger than $l_0 + \frac{\mu mg}{k}$ before the terminating peak (which may be the first peak) while x-coordinates of the troughs will always be smaller than $l_0 - \frac{\mu mg}{k}$ before the terminating trough. This implies that the conditions for stopping at the $(n+1)$th peak and trough are $x'_n \leq l_0 + \frac{\mu mg}{k}$ and $x_n \geq l_0 - \frac{\mu mg}{k}$. Thus, the minimum n's for the motion to stop at the $(n+1)$th peak and trough are respectively

$$n_{peak} = \left\lceil \frac{l_0 k}{4\mu mg} - \frac{3}{4} \right\rceil,$$
$$n_{trough} = \left\lceil \frac{l_0 k}{4\mu mg} - \frac{1}{4} \right\rceil.$$

The total number of cycles completed by m is the minimum of the two values, $\min(n_{peak}, n_{trough})$.

17. Stabilizer**

Since this problem is purely one-dimensional, define the x-axis to be along the direction of concern. Let the x-coordinates of m_1 and m_2 be x_1 and x_2 respectively. Since the relaxed length of the spring does not affect the oscillation, we can simply take it to be zero. Equivalently, we could have defined x_1 and x_2 to be the displacements of the respective masses from their equilibrium position in the absence of the external driving force. Writing their equations of motion,

$$m_1\ddot{x}_1 = f\cos\omega t + k(x_2 - x_1),$$
$$m_2\ddot{x}_2 = -k(x_2 - x_1).$$

Multiplying the second equation by m_1 and subtracting it by the first equation multiplied by m_2,

$$m_1m_2(\ddot{x}_2 - \ddot{x}_1) = -k(m_1 + m_2)(x_2 - x_1) - m_2 f\cos\omega t.$$

Using the substitution $u = x_2 - x_1$,

$$\ddot{u} + \frac{k(m_1 + m_2)}{m_1m_2}u = -\frac{f}{m_1}\cos\omega t.$$

To obtain the particular solution to the above equation, we can solve the following differential equation and take the real component of its particular solution.

$$\ddot{u} + \frac{k(m_1 + m_2)}{m_1m_2}u = -\frac{f}{m_1}e^{i\omega t}.$$

Substituting the trial solution $u = Ae^{i\omega t}$,

$$-A\omega^2 + \frac{k(m_1 + m_2)}{m_1m_2}A = -\frac{f}{m_1}$$

$$A = \frac{m_2 f}{m_1m_2\omega^2 - k(m_1 + m_2)}.$$

The particular solution to our original equation is obtained by taking the real component of $u = Ae^{i\omega t}$, that is

$$u = A\cos\omega t.$$

Substituting this expression of u into $m_1\ddot{x}_1 = f\cos\omega t + ku$ yields

$$\ddot{x}_1 = \frac{f}{m_1}\cos\omega t + \frac{km_2 f}{m_1[m_1m_2\omega^2 - k(m_1 + m_2)]}\cos\omega t.$$

These terms cancel when

$$k = \omega^2 m_2.$$

18. Masses on Hoop***

Label the masses m, $2m$ and $2m$ from one to three, in a clockwise fashion around the hoop. Then, let θ_1, θ_2 and θ_3 represent the angular displacements from their equilibrium positions. Let $x_i = R\theta_i$ for the sake of convenience. Then, the equation of motions of the three masses can be shown to be

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{pmatrix} = \begin{pmatrix} -\frac{4k}{m} & \frac{2k}{m} & \frac{2k}{m} \\ \frac{k}{m} & -\frac{3k}{2m} & \frac{k}{2m} \\ \frac{k}{m} & \frac{k}{2m} & -\frac{3k}{2m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Guessing solutions of the form $\boldsymbol{X} = \boldsymbol{u}e^{i\omega t}$,

$$\begin{pmatrix} -\frac{4k}{m} + \omega^2 & \frac{2k}{m} & \frac{2k}{m} \\ \frac{k}{m} & -\frac{3k}{2m} + \omega^2 & \frac{k}{2m} \\ \frac{k}{m} & \frac{k}{2m} & -\frac{3k}{2m} + \omega^2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 0.$$

For non-trivial solutions to exist, the determinant of the first matrix must be zero. One can show after some simplification that this is equivalent to

$$\omega^2 \left(\omega^2 - \frac{2k}{m}\right) \left(\omega^2 - \frac{5k}{m}\right) = 0.$$

Therefore, the normal frequencies are $\omega_1 = \sqrt{\frac{2k}{m}}$ and $\omega_2 = \sqrt{\frac{5k}{m}}$. Now, we shall determine the corresponding normal modes. Substituting $\omega_1^2 = \frac{2k}{m}$ into the matrix equation,

$$\begin{pmatrix} -\frac{2k}{m} & \frac{2k}{m} & \frac{2k}{m} \\ \frac{k}{m} & \frac{k}{2m} & \frac{k}{2m} \\ \frac{k}{m} & \frac{k}{2m} & \frac{k}{2m} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 0.$$

Solving for $\boldsymbol{u}$, the eigenspace associated with the eigenvalue $\lambda_1 = -\omega_1^2 = -\frac{2k}{m}$ is the collection of vectors

$$\boldsymbol{E}_{\frac{2k}{m}} = C_1 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

for some scalar C_1. Therefore, the normal mode associated with $\omega_1 = \sqrt{\frac{2k}{m}}$ is $(0, -1, 1)$. Substituting $\omega_2^2 = \frac{5k}{m}$ into the matrix equation,

$$\begin{pmatrix} \frac{k}{m} & \frac{2k}{m} & \frac{2k}{m} \\ \frac{k}{m} & \frac{7k}{2m} & \frac{k}{2m} \\ \frac{k}{m} & \frac{k}{2m} & \frac{7k}{2m} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = 0.$$

The eigenspace associated with the eigenvalue $\lambda_2 = -\omega_2^2 = -\frac{5k}{m}$ is

$$\boldsymbol{E}_{\frac{5k}{m}} = C_2 \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix}$$

for some scalar C_2. Therefore, the only normal mode associated with $\omega_2 = \sqrt{\frac{5k}{m}}$ is $(-4, 1, 1)$. Lastly, notice that we only have four parameters from these two solutions. Therefore, another independent solution is needed for the general solution. Notice that $\omega^2 = 0$ also causes the determinant to be zero — suggesting that the second derivatives of the displacements are zero. Then, it is wise to guess a solution of the form $\boldsymbol{u}(C_3 t + C_4)$ where C_3 and C_4 are real. Substituting this into the original matrix equation would yield vectors of the form

$$\boldsymbol{E}_0 = C_5 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

as solutions for $\boldsymbol{u}$ in such cases. This makes sense as it just means that the masses are displaced by the same amount — causing the lengths of the springs to remain constant. The general linear combination of the solutions is

$$\boldsymbol{E}_{\frac{2k}{m}} e^{i\sqrt{\frac{2k}{m}}t} + \boldsymbol{E}_{\frac{5k}{m}} e^{i\sqrt{\frac{5k}{m}}t} + \boldsymbol{E}_0 (C_3 t + C_4).$$

The displacements of the masses are obtained by taking the real component of the above.

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = D_1 \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} \cos\left(\sqrt{\frac{2k}{m}}t + \phi_1\right)$$
$$+ D_2 \begin{pmatrix} -4 \\ 1 \\ 1 \end{pmatrix} \cos\left(\sqrt{\frac{5k}{m}}t + \phi_2\right) + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (D_3 t + D_4),$$

where D_1, D_2, D_3 and D_4 are real scalars.

Chapter 11

Non-Inertial Frames

Even though Newton's laws are only valid in inertial frames, certain modifications can be made to extend them to accelerating and rotating frames as well. This enables us to furnish a more accurate portrayal of systems on Earth — an inherently non-inertial frame that is rotating about its axis and revolving around the Sun.

11.1 Purely Accelerating Frame

Consider a train undergoing a possibly time-varying acceleration $\ddot{\boldsymbol{R}}$ with respect to the ground, which is presumed to be an inertial frame. How should the equations of motion of an experiment conducted in the train look like in the train's frame? Intuitively, one would expect that particles will experience an additional $-\ddot{\boldsymbol{R}}$ acceleration superimposed on the acceleration they would have experienced in the ground frame — this is tantamount to each particle experiencing an additional fictitious $-m\ddot{\boldsymbol{R}}$ force in the train's frame, where m is its mass.

Let us take a more formal approach to this. The core idea behind deriving the equations of motion in non-inertial frames is to apply Newton's laws in an inertial frame and then express the coordinates of the inertial frame in terms of the coordinates of the non-inertial frame.

Let $\boldsymbol{r}_0$ be the position vector of a particular particle of interest with respect to the ground frame and $\boldsymbol{r}$ be that with respect to the train's frame. Let $\boldsymbol{R}$ denote the origin of the train's frame with respect to the ground frame such that

$$\boldsymbol{r}_0 = \boldsymbol{R} + \boldsymbol{r} \tag{11.1}$$

$$\implies \ddot{\boldsymbol{r}}_0 = \ddot{\boldsymbol{R}} + \ddot{\boldsymbol{r}}. \tag{11.2}$$

From Newton's second law, we know that the net external force $\sum \boldsymbol{F}$ on a particle induces a proportional acceleration in an inertial frame.

$$\sum \boldsymbol{F} = m\ddot{\boldsymbol{r}}_0, \tag{11.3}$$

where m is the mass of the particle. Multiplying Eq. (11.2) by m and substituting Eq. (11.3),

$$\sum \boldsymbol{F} = m\ddot{\boldsymbol{R}} + m\ddot{\boldsymbol{r}}.$$

Rearranging,

$$m\ddot{\boldsymbol{r}} = \sum \boldsymbol{F} - m\ddot{\boldsymbol{R}}. \tag{11.4}$$

It can be seen that the equation of motion of a particle in the train's frame is akin to Newton's law, provided that we introduce a fictitious $-m\ddot{\boldsymbol{R}}$ force known as the inertial force.

$$\boldsymbol{F}_{inertial} = -m\ddot{\boldsymbol{R}}. \tag{11.5}$$

As a word of caution, $\sum \boldsymbol{F}$ is the sum of the real forces which are invariant across different frames. However, $\boldsymbol{F}_{inertial}$ is not a physical force and varies across different frames as it originates from the changing origin of the accelerating frame rather than interactions with concrete entities.

The inertial force is ubiquitous in real life. When the bus you take suddenly accelerates forward at $\ddot{\boldsymbol{R}}$, an external ground observer would remark that the friction force on your feet must be $m\ddot{\boldsymbol{R}}$ where m is your mass for you to remain stationary with respect to the train. In the train's frame (which is also your frame in this case, as you remain relatively still), you would explain that the friction force $m\ddot{\boldsymbol{R}}$ balances the inertial force $-m\ddot{\boldsymbol{R}}$ such that you do not accelerate in this frame. On the other hand, if there were no friction, the inertial force would simply fling you towards the back of the train until you are cushioned by a barricade.

Problem: When a train with a frictionless floor accelerates forwards uniformly, a ball on the floor is flung towards the back. Now, consider a light balloon filled with helium that is suspended from the ceiling of the train. At equilibrium, in what direction does the balloon tilt towards?

In the train's frame, there is an inertial force on the enclosed particles, directed towards the back. When air in the train eventually attains equilibrium, the pressure must decrease from the back to the front of the train to counteract the inertial force on each air section. This pressure gradient then results in a force on the balloon whose horizontal component is directed towards the front of the train. Since the inertial force on the balloon is negligible due to its minuscule mass, the balloon must tilt towards the front of the train so that the horizontal component of the constraint force of the ceiling on the balloon can be directed towards the back of the train to balance the forward force due to the tapering pressure.

Rigid Bodies

It is not hard to see that the total inertial force on a rigid body with mass m should be $-m\ddot{\boldsymbol{R}}$ as each mass element dm experiences an inertial force $-dm\ddot{\boldsymbol{R}}$ which symbolizes that the total force is $-\int \ddot{\boldsymbol{R}} dm = -m\ddot{\boldsymbol{R}}$. However, where should the effective inertial force $-m\ddot{\boldsymbol{R}}$ act at? The intuitive approach is to notice the analogy between the inertial force and a region of uniform gravitational field $-\ddot{\boldsymbol{R}}$. The object cannot tell if it is under the influence of the former or latter — implying that the effective inertial force $-m\ddot{\boldsymbol{R}}$ should act at the center of mass of the object, in a manner similar to the effective gravitational force. More rigorously, we can show that the torque produced by the inertial force on a rigid body about an arbitrary origin is equivalent to that due to an effective inertial force $-m\ddot{\boldsymbol{R}}$ on the center of mass of the object, about the same origin. Incidentally, this is also the formal proof behind the fact that the effective gravitational force due to a uniform field acts at the center of mass. Let $\boldsymbol{r}$ denote the position vector of an infinitesimal mass element dm with respect to an arbitrary origin. Since the torque on this element is $-\boldsymbol{r} \times \ddot{\boldsymbol{R}} dm$, the total torque on the extended body of interest is

$$\int -\boldsymbol{r} \times \ddot{\boldsymbol{R}} dm = \left(\int \boldsymbol{r} dm \right) \times -\ddot{\boldsymbol{R}} = m\boldsymbol{r}_{CM} \times -\ddot{\boldsymbol{R}} = \boldsymbol{r}_{CM} \times (-m\ddot{\boldsymbol{R}}),$$

as $\int \boldsymbol{r} dm = m\boldsymbol{r}_{CM}$ where $\boldsymbol{r}_{CM}$ is the position vector of the center of mass by definition. The above shows that the effective inertial force $-m\ddot{\boldsymbol{R}}$ resides at the center of mass, $\boldsymbol{r}_{CM}$.

Problem: A uniform equilateral triangle of length l rests on the rough base of a truck. The truck then undergoes a uniform acceleration a rightwards (in the plane of the triangle). Assuming that the coefficient of static friction μ is large enough, determine the largest acceleration a_{max} for which the triangle will remain static. What is the minimum μ for which the triangle remains static for all accelerations smaller than a_{max}? Finally, show that if $a > a_{max}$ and if the triangle subsequently rotates about its left vertex (which remains stationary relative to the ground), the triangle inevitably topples over.

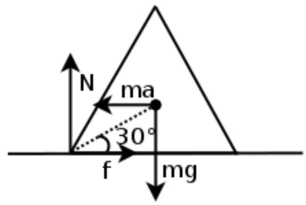


Figure 11.1: Forces on triangle

Switch to the accelerating frame of the truck and assume that the triangle is static. The forces on the triangle are the inertial force, its weight, the normal force due to the ground and friction. Balancing torques about the left vertex, the anti-clockwise torques due to the inertial force (which acts at the center of mass) and the normal force must nullify the clockwise torque due to the triangle's weight. Generally, the normal force will be distributed along the base of the triangle and act in the upwards direction — producing an anti-clockwise torque that is non-trivial to calculate. However, in the boundary case where a is maximum, the anti-clockwise torque due to the normal force must be minimum — implying that it should reside exactly at the left vertex and contribute zero torque. Then, the torques due to the inertial force and the weight of the triangle must cancel each other — implying that the net force vector due to both forces, emanating from the center of mass, must pass through the left vertex. In such a situation,

$$\frac{g}{a_{max}} = \tan 30^\circ = \frac{1}{\sqrt{3}}$$

$$a_{max} = \sqrt{3}g.$$

The inertial force in this case is ma_{max} leftwards, where m is the mass of the triangle, which implies that the friction force must be $f = ma_{max}$ rightwards to ensure translational equilibrium in the horizontal direction. Since the normal force must constantly be mg to balance the weight of the object, the coefficient of static friction μ must satisfy

$$\mu \geq \frac{f}{N} = \frac{a_{max}}{g} = \sqrt{3}$$

for the triangle to remain static for all accelerations smaller than a_{max}. Finally, when $a > a_{max}$, the triangle can no longer stay stationary and begins to rotate about its left vertex. To show that the triangle eventually topples over, we simply have to prove that the triangle is able to attain the configuration where a vertical line along its center of mass passes through the left vertex that it rotates about. To this end, we simply have to show that the kinetic energy of the triangle is positive for all configurations before this particular configuration. Now, one might wonder if the angular velocity of the triangle might be in the wrong direction (tending to turn it back into the original position) while its kinetic energy is still positive. The resolution to this query is that there cannot be a sudden discontinuity in the sign of the triangle's angular velocity (as the torque is finite) — if it begins at a positive value, it will remain positive until it attains a null value (which we shall prove to be impossible when it is in the process of rotating).

Instead of computing the kinetic energy of the triangle directly, we can make the astute observation that the inertial force is conservative such that we can ascribe a potential energy function to it (similar to the gravitational potential energy with $-\boldsymbol{a}$ being the uniform gravitational field where $\boldsymbol{a}$ is the acceleration of the truck) and demand that the change in the total potential energy of the triangle is negative to show that its kinetic energy is positive (since its initial kinetic energy is zero). Define the origin at the fixed vertex that the triangle rotates about and the x and y axes to be positive rightwards and upwards respectively. The potential energy of the triangle due to the inertial force is max where x is the x-coordinate of its center of mass. Denoting l as the length of the segment connecting the origin to the center of mass and θ as the angle subtended by the position vector of the center of mass and the vertical,

$$x = l \sin\theta.$$

The total potential energy of the triangle is

$$U(\theta) = mal \sin\theta + mgl \cos\theta.$$

The initial potential energy is

$$U\left(60^\circ\right) = \frac{\sqrt{3}}{2} mal + \frac{1}{2} mgl.$$

The change in potential energy from the initial state to a state at angle θ is

$$\Delta U = mal\left(\sin\theta - \frac{\sqrt{3}}{2}\right) + mgl\left(\cos\theta - \frac{1}{2}\right).$$

For the sake of convenience, express a as $a = kg$ for some constant k, so that

$$\Delta U = mgl\left(k\sin\theta - \frac{k\sqrt{3}}{2} + \cos\theta - \frac{1}{2}\right).$$

Our objective is to show that this is negative for all $0 < \theta < 60°$ — this is equivalent to showing that

$$k \sin\theta - \frac{k\sqrt{3}}{2} + \cos\theta - \frac{1}{2} < 0.$$

Substituting our boundary case $a = a_{max}$ such that $k = \sqrt{3}$, we wish to prove

$$\sqrt{3}\sin\theta - \frac{3}{2} + \cos\theta - \frac{1}{2} < 0$$
$$\implies 2\cos(\theta - 60°) - 2 < 0$$

for $0 < \theta < 60°$. This is evidently true as $\cos(\theta - 60°) < 1$ for $0 < \theta < 60°$ (the maximum value of 1 occurs at the prohibited value of $\theta = 60°$). Finally, to show that the relevant inequality is valid for larger values of k, simply observe that $\sin\theta - \frac{\sqrt{3}}{2} < 0$ for $0 < \theta < 60°$ such that a larger value of k only exacerbates the negative value of the left-hand side (through the $k\sin\theta - k\frac{\sqrt{3}}{2}$ term).

11.1.1 *Tides*

A classic application of the inertial force pertains to the rises and falls of the sea levels on Earth, termed as high and low tides respectively. Tides arise due to the disparity in the acceleration of different parts on Earth due to the gravitational forces of the Sun and the Moon. Though the land on Earth can be approximated as a spherical rigid body, the oceans are free to flow around and thus can undertake non-spherical shapes due to the varying accelerations at different points on Earth. A common misconception is that the Sun or Moon pulls water towards the closer side of the Earth — forming a single bulge. However, the paramount effect here is the difference in accelerations. It is true that a mass on the closer side of the Earth (relative to the Sun or Moon) experiences a greater acceleration than Earth, but it is also true that the Earth experiences a greater acceleration than a mass located at the further end of the Earth. The latter factor causes another bulge to develop at the further end of the Earth — explaining the two observed high tides per day. Ultimately, the formation of tides is a relative phenomenon.

Let us analyze the effect of the gravity of a single spherical object, of uniform mass M, on Earth first. Define $\boldsymbol{r}_0$ as the position vector of a mass m located on the surface of Earth with respect to the center of the massive object, $\boldsymbol{r}$, as that with respect to the center of Earth and $\boldsymbol{R}$ as the position

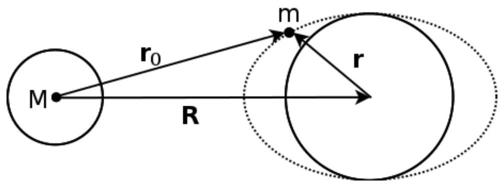


Figure 11.2: Distortion of oceans due to massive object

vector of the Earth with respect to the center of the massive object. Since the Earth accelerates at $-\frac{GM}{R^3}\boldsymbol{R}$, the inertial force on m in the Earth's frame is

$$\boldsymbol{F}_{inertial} = \frac{GMm}{R^3}\boldsymbol{R}.$$

The forces on m (including the fictitious forces) in the Earth's frame can be divided into three groups — the gravitational force due to M, the inertial force $\boldsymbol{F}_{inertial}$ and all other forces denoted by $\sum \boldsymbol{F}_{others}$ (e.g. the gravitational force due to Earth).

$$\boldsymbol{F}_{net} = -\frac{GMm}{r_0^3}\boldsymbol{r}_0 + \frac{GMm}{R^3}\boldsymbol{R} + \sum \boldsymbol{F}_{others}.$$

The last term persists in the absence of M while the first two terms arise from the presence of M — the combination of these two additional terms is known as the tidal force $\boldsymbol{F}_{tidal}$, where

$$\boldsymbol{F}_{tidal} = -GMm\left(\frac{\boldsymbol{r}_0}{r_0^3} - \frac{\boldsymbol{R}}{R^3}\right). \tag{11.6}$$

This tidal force is the impetus behind the deformation of the oceans. Usually, $r \ll R$ so we can make a few approximations and discard second-order terms in $\frac{r}{R}$ to simplify the expression for the tidal force. Since $\boldsymbol{r}_0 = \boldsymbol{R} + \boldsymbol{r}$,

$$\boldsymbol{F}_{tidal} = -GMm\left(\frac{\boldsymbol{R}+\boldsymbol{r}}{|\boldsymbol{R}+\boldsymbol{r}|^3} - \frac{\boldsymbol{R}}{R^3}\right).$$

The magnitude of $\boldsymbol{r}_0$ can be simplified via a binomial expansion.

$$|\boldsymbol{R}+\boldsymbol{r}| = \sqrt{R^2 + r^2 + 2\boldsymbol{R}\cdot\boldsymbol{r}} \approx R\sqrt{1 + \frac{2\boldsymbol{R}\cdot\boldsymbol{r}}{R^2}} \approx R\left(1 + \frac{\boldsymbol{R}\cdot\boldsymbol{r}}{R^2}\right),$$

$$\boldsymbol{F}_{tidal} = -GMm\left(\frac{\boldsymbol{R}+\boldsymbol{r}}{R^3\left(1+\frac{\boldsymbol{R}\cdot\boldsymbol{r}}{R^2}\right)^3} - \frac{\boldsymbol{R}}{R^3}\right) \tag{11.7}$$

$$\approx -GMm\left(\frac{\boldsymbol{R}+\boldsymbol{r}}{R^3}\left(1-3\frac{\boldsymbol{R}\cdot\boldsymbol{r}}{R^2}\right)-\frac{\boldsymbol{R}}{R^3}\right) \tag{11.8}$$

$$\approx -\frac{GMm}{R^3}(\boldsymbol{r}-3(\hat{\boldsymbol{R}}\cdot\boldsymbol{r})\hat{\boldsymbol{R}}). \tag{11.9}$$

Let us consider a few special cases of the tidal force to get an intuitive feeling for the shape of the ocean. When $\hat{\boldsymbol{R}}$ is parallel to $\boldsymbol{r}$ (i.e. at the closest and furthest ends), $(\hat{\boldsymbol{R}}\cdot\boldsymbol{r})\hat{\boldsymbol{R}}=\boldsymbol{r}$.

$$\boldsymbol{F}_{tidal}=\frac{2GMm\boldsymbol{r}}{R^3}.$$

That is, the tidal force tends to push water radially outwards at these locations — resulting in high tides of equal magnitude at the left and right ends. When $\hat{\boldsymbol{R}}$ is perpendicular to $\boldsymbol{r}$ at the top and bottom of Earth in the plane containing m and the centers of M and the Earth, $\hat{\boldsymbol{R}}\cdot\boldsymbol{r}=0$.

$$\boldsymbol{F}_{tidal}=-\frac{GMm\boldsymbol{r}}{R^3}.$$

The tidal force "attracts" water towards the center of the Earth — resulting in low tides at these locations. With a rough gauge of how the ocean surface should look like, we can actually calculate the height of the high tides with this model. The trenchant observation here is that the other forces on a piece of fluid $\sum \boldsymbol{F}_{others}$ consist of $m\boldsymbol{g}$, the gravitational force due to Earth, and the buoyant force due to neighboring fluid segments which is perpendicular to the surface of the fluid, as a fluid cannot withstand or exert any shear force (forces parallel to its surface) without deformation. Therefore, for a piece of fluid to not deform, the gravitational force due to the Earth and the tidal force must be perpendicular to its surface. If we are able to define a potential energy function for the tidal force, the surface of the ocean must correspond to an equipotential surface as the force, which is the negative gradient of the potential energy, must be perpendicular to the surface.

It is easy to guess a potential energy function for the tidal force if we express Eq. (11.9) in terms of Cartesian coordinates. Define the x and y axes to be positive rightwards and upwards, with the origin located at the center of Earth. If an element of mass m is located at (x, y),

$$\boldsymbol{F}_{tidal}=\frac{2GMmx}{R^3}\hat{\boldsymbol{\imath}}-\frac{GMmy}{R^3}\hat{\boldsymbol{\jmath}}.$$

One can check that the negative gradient[1] of

$$U_{tidal} = -\frac{GMmx^2}{R^3} + \frac{GMmy^2}{2R^3} \tag{11.10}$$

indeed results in the force above — implying that it is the correct tidal potential energy. The total potential energy of a fluid element is the sum of the tidal potential energy and the gravitational potential energy associated with its interactions with Earth. The latter shall be denoted by U_{grav}. Since the ocean surface is equipotential, the total potential energy at the top end should be equal to that at the right end, and

$$U_{grav,top} + U_{tidal,top} = U_{grav,right} + U_{tidal,right},$$
$$U_{grav,right} - U_{grav,top} = U_{tidal,top} - U_{tidal,right}.$$

The left-hand side is simply mgh where g is the gravitational field strength at the surface of Earth and h is the height difference between the low and high tides as h should be small when compared to R_e, the radius of Earth. In fact, h can also be taken to be the altitude of the high tide as the altitude of the low tide should be approximately zero. In evaluating the tidal energies, the coordinates of the top and right ends can be taken at $(0, R_e)$ and $(R_e, 0)$ respectively as h is small compared to R_e, and is hence even smaller when compared to R. Substituting these expressions,

$$mgh = \frac{3GMmR_e^2}{2R^3}$$
$$h = \frac{3GMR_e^2}{2gR^3}.$$

By Gauss' law, $g = \frac{GM_e}{R_e^2}$ where M_e is the mass of Earth. Thus,

$$h = \frac{3MR_e^4}{2M_eR^3}. \tag{11.11}$$

Using the actual parameters ($M_e = 5.98 \times 10^{24}$kg, $R_e = 6.37 \times 10^6$m for Earth and $M = 7.35 \times 10^{22}$kg, $R = 3.84 \times 10^8$m for the Moon), the height

[1] Actually, we can also easily guess a potential energy function for the exact tidal force given by Eq. (11.6) but subsequent approximations would still yield the same expression as Eq. (11.10).

of the high tide caused individually by the Moon is

$$h_{Moon} = 54\text{cm}.$$

On the other hand, using the values ($M = 1.99 \times 10^{30}$kg and $R = 1.50 \times 10^{11}$m) for the Sun yields

$$h_{Sun} = 25\text{cm}.$$

Since both of these values are significant, the height of high tides depends on the relative orientations of the Earth, Moon and Sun. If the centers are collinear, the effects of the Moon and Sun are reinforced such that the predicted value for the high tide attains the maximum altitude

$$h_{spring} = 54 + 25 = 79\text{cm}.$$

These tides are known as spring tides. If the lines joining the center of Earth to the Moon and Sun are mutually perpendicular, the effects of the Moon and Sun counteract each other such that high tides of the minimum altitude, known as neap tides, are formed:

$$h_{neap} = 54 - 25 = 29\text{cm}.$$

That said, take these altitudes with a pinch of salt as there are many complications that this model has not accounted for. For example, the Earth is not perfectly spherical due in part to the equatorial bulge stemming from the centrifugal force (discussed later). The existence of bordering continents in certain regions also help to clump water together, producing larger tides. However, the order of magnitude of these altitudes are consistent with the observed values, so the above model is still useful in this sense.

11.2 Accelerating and Rotating Frame

Most generally, a frame can accelerate translationally and rotate at an angular velocity $\boldsymbol{\omega}$ with respect to an inertial frame. It was shown in Section 3.5.1 that the rate of change of a vector $\boldsymbol{A}$ of fixed length emanating from a fixed origin and rotating at an angular velocity $\boldsymbol{\omega}$ in a frame is

$$\frac{d\boldsymbol{A}}{dt} = \boldsymbol{\omega} \times \boldsymbol{A}. \tag{11.12}$$

This relationship will be very useful later. Adopting the same notation as the previous section, let $\boldsymbol{r}_0$ be the position vector of a particle of interest with respect to an inertial frame S, $\boldsymbol{r}$ be that with respect to a non-inertial frame S' of concern and $\boldsymbol{R}$ be the position vector of the origin O$'$ of S' with

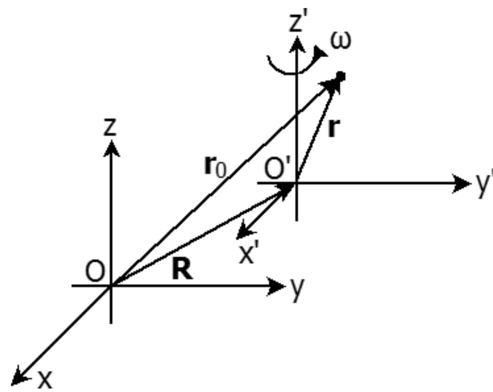


Figure 11.3: Inertial frame S and non-inertial frame S'

respect to the origin O of S (Fig. 11.3). S' possesses an angular velocity $\boldsymbol{\omega}$ and possibly accelerates translationally relative to S.

From the basic principles of vector addition,

$$\boldsymbol{r}_0 = \boldsymbol{R} + \boldsymbol{r} \implies \frac{d^2\boldsymbol{r}_0}{dt^2} = \frac{d^2\boldsymbol{R}}{dt^2} + \frac{d^2\boldsymbol{r}}{dt^2}.$$

We will stick to the notation $\frac{d}{dt}$ for time derivatives instead of the dot notation, for the sake of greater clarity. From Newton's second law, we know that

$$\sum \boldsymbol{F} = m\frac{d^2\boldsymbol{r}_0}{dt^2},$$

where m is the mass of the particle. The two equations above yield

$$\sum \boldsymbol{F} = m\frac{d^2\boldsymbol{R}}{dt^2} + m\frac{d^2\boldsymbol{r}}{dt^2}. \tag{11.13}$$

Our objective is to relate the acceleration as perceived in the non-inertial frame S', to $\sum \boldsymbol{F}$. To this end, though the equation above quintessentially involves vector quantities which can be evaluated in any frame and are frame-independent, we should express $\boldsymbol{r}$ in terms of the basis vectors $\hat{\boldsymbol{i}}'$, $\hat{\boldsymbol{j}}'$ and $\hat{\boldsymbol{k}}'$ in S', as

$$\boldsymbol{r} = r_{x'}\hat{\boldsymbol{i}}' + r_{y'}\hat{\boldsymbol{j}}' + r_{z'}\hat{\boldsymbol{k}}'.$$

Instead of deriving $\frac{d^2\boldsymbol{r}}{dt^2}$ directly, we can first create an important tool that will expedite this process. Let us determine $\frac{d\boldsymbol{A}}{dt}$ for a general vector $\boldsymbol{A}$ given by

$$\boldsymbol{A} = A_{x'}\hat{\boldsymbol{i}}' + A_{y'}\hat{\boldsymbol{j}}' + A_{z'}\hat{\boldsymbol{k}}'.$$

The rate of change of $\boldsymbol{A}$ emerges from the changes in the components of $\boldsymbol{A}$ in the coordinate system of S' and the change in the basis vectors of S':

$$\frac{d\boldsymbol{A}}{dt} = \left(\frac{dA_{x'}}{dt}\hat{\boldsymbol{i}}' + \frac{dA_{y'}}{dt}\hat{\boldsymbol{j}}' + \frac{dA_{z'}}{dt}\hat{\boldsymbol{k}}'\right) + A_{x'}\frac{d\hat{\boldsymbol{i}}'}{dt} + A_{y'}\frac{d\hat{\boldsymbol{j}}'}{dt} + A_{z'}\frac{d\hat{\boldsymbol{k}}'}{dt}.$$

The first three terms correspond to the rate of change of $\boldsymbol{A}$ as observed in the non-inertial frame, as the basis vectors of a particular frame are not perceived to change in that frame. We will denote them as $\frac{d\boldsymbol{A}}{dt}\big|_{rot}$. Next, to determine the rate of change of the basis vectors, notice that these vectors are of fixed length and are rotating at angular velocity $\boldsymbol{\omega}$ with respect to S. Applying Eq. (11.12),

$$\frac{d\hat{\boldsymbol{i}}'}{dt} = \boldsymbol{\omega} \times \hat{\boldsymbol{i}}',$$

$$\frac{d\hat{\boldsymbol{j}}'}{dt} = \boldsymbol{\omega} \times \hat{\boldsymbol{j}}',$$

$$\frac{d\hat{\boldsymbol{k}}'}{dt} = \boldsymbol{\omega} \times \hat{\boldsymbol{k}}'.$$

Overall,

$$\frac{d\boldsymbol{A}}{dt} = \frac{d\boldsymbol{A}}{dt}\Big|_{rot} + \boldsymbol{\omega} \times \left(A_{x'}\hat{\boldsymbol{i}}' + A_{y'}\hat{\boldsymbol{j}}' + A_{z'}\hat{\boldsymbol{k}}'\right)$$

$$\frac{d\boldsymbol{A}}{dt} = \frac{d\boldsymbol{A}}{dt}\Big|_{rot} + \boldsymbol{\omega} \times \boldsymbol{A}. \tag{11.14}$$

We can express the above in a more illuminating form, in terms of operators.

$$\frac{d}{dt} = \frac{d}{dt}\Big|_{rot} + \boldsymbol{\omega}\times$$

When we apply $\frac{d}{dt}$ to a vector, it is equivalent to applying the right-hand side to that vector too. Armed with this tool, we can compute $\frac{d\boldsymbol{r}}{dt}$ in our original problem by applying the above to $\boldsymbol{r}$.

$$\frac{d\boldsymbol{r}}{dt} = \frac{d\boldsymbol{r}}{dt}\Big|_{rot} + \boldsymbol{\omega} \times \boldsymbol{r}. \tag{11.15}$$

Applying the operator again, we can procure our desired $\frac{d^2\boldsymbol{r}}{dt^2}\big|_{rot}$ (acceleration as observed in S').

$$\begin{aligned}
\frac{d^2\boldsymbol{r}}{dt^2} &= \frac{d}{dt}\left(\frac{d\boldsymbol{r}}{dt}\Big|_{rot}\right) + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r} + \boldsymbol{\omega} \times \frac{d\boldsymbol{r}}{dt} \\
&= \left(\frac{d}{dt}\Big|_{rot} + \boldsymbol{\omega}\times\right)\frac{d\boldsymbol{r}}{dt}\Big|_{rot} + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r} + \boldsymbol{\omega} \times \left(\frac{d\boldsymbol{r}}{dt}\Big|_{rot} + \boldsymbol{\omega} \times \boldsymbol{r}\right)
\end{aligned}$$

$$= \frac{d^2\boldsymbol{r}}{dt^2}\Big|_{rot} + \boldsymbol{\omega} \times \frac{d\boldsymbol{r}}{dt}\Big|_{rot} + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r} + \boldsymbol{\omega} \times \frac{d\boldsymbol{r}}{dt}\Big|_{rot} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})$$

$$= \frac{d^2\boldsymbol{r}}{dt^2}\Big|_{rot} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + 2\boldsymbol{\omega} \times \frac{d\boldsymbol{r}}{dt}\Big|_{rot} + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r}.$$

For the sake of brevity, denote $\boldsymbol{a}_{rot} = \frac{d^2\boldsymbol{r}}{dt^2}\big|_{rot}$ and $\boldsymbol{v}_{rot} = \frac{d\boldsymbol{r}}{dt}\big|_{rot}$ as the acceleration and velocity observed in the non-inertial frame, respectively. Then,

$$\frac{d^2\boldsymbol{r}}{dt} = \boldsymbol{a}_{rot} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + 2\boldsymbol{\omega} \times \boldsymbol{v}_{rot} + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r}. \tag{11.16}$$

Substituting Eq. (11.16) into (11.13),

$$\sum \boldsymbol{F} = m\left(\frac{d^2\boldsymbol{R}}{dt^2} + \boldsymbol{a}_{rot} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + 2\boldsymbol{\omega} \times \boldsymbol{v}_{rot} + \frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r}\right).$$

Rearranging,

$$m\boldsymbol{a}_{rot} = \sum \boldsymbol{F} - m\frac{d^2\boldsymbol{R}}{dt^2} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + 2m\boldsymbol{v}_{rot} \times \boldsymbol{\omega} - m\frac{d\boldsymbol{\omega}}{dt} \times \boldsymbol{r} \tag{11.17}$$

$$= \sum \boldsymbol{F} + \boldsymbol{F}_{inertial} + \boldsymbol{F}_{cen} + \boldsymbol{F}_{cor} + \boldsymbol{F}_{azi}, \tag{11.18}$$

where the corresponding fictitious forces are termed the inertial, centrifugal, Coriolis and azimuthal forces respectively. Since the inertial force has already been explicated, we proceed with the centrifugal force.

11.3 Centrifugal Force

The origin of the centrifugal force $\boldsymbol{F}_{cen} = -m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r})$ becomes lucid when we first view the situation from an inertial frame. Consider a ball that is connected to a fixed pivot via an inextensible string of length r on a frictionless table. If the ball undergoes circular motion with an instantaneous angular speed ω, the instantaneous tension in the string must be $-mr\omega^2$, where the negative sign indicates that the force is directed radially inwards, to provide the required centripetal acceleration. On the other hand, if we switch to the frame of the ball, the centrifugal force on the ball is

$$\boldsymbol{F}_{cen} = -m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) = -m\left[\boldsymbol{\omega}(\boldsymbol{\omega} \cdot \boldsymbol{r}) - \boldsymbol{r}\left(\boldsymbol{\omega} \cdot \boldsymbol{\omega}\right)\right] = 0 + m\omega^2\boldsymbol{r}, \tag{11.19}$$

as $\boldsymbol{\omega}$ is perpendicular to $\boldsymbol{r}$ in this case. As implied by its nomenclature, the centrifugal force is directed radially outwards as it tends to throw objects away from the center. We say that in this frame, the tension on the ball balances the centrifugal force. In a certain sense, the cause and effect are reversed. In an inertial frame, some force must provide the centripetal acceleration required by a particle to undergo circular motion, while in a frame rotating at $\boldsymbol{\omega}$ relative to the inertial frame, the centrifugal force acts as a litmus test for whether the particle can indeed rotate at $\boldsymbol{\omega}$ in the inertial frame by checking if some force can counteract it. Note that these statements easily extend to the more general case where $\boldsymbol{\omega}$ is not perpendicular to $\boldsymbol{r}$ — the centrifugal force is simply directed radially outwards from the center of rotation in the plane perpendicular to ω that contains the rotating particle (which is at a perpendicular distance $r\sin\theta$ from the rotational axis $\boldsymbol{\omega}$, where θ is the angle between $\boldsymbol{\omega}$ and $\boldsymbol{r}$).

Problem: Modeling the Earth as a uniform and rotating sphere, explain why a plumb line (the string connecting a bob to a pivot at equilibrium) on Earth does not point towards the center of Earth in general. Determine the angle α that the plumb line makes with the true gravitational force on the bob due to Earth as a function of θ, the colatitude on Earth at which the experiment is conducted, in terms of self-defined parameters. Note that the colatitude on a point on the surface of Earth refers to the angle subtended by the line joining the center of the Earth to the North pole and the line joining the center of the Earth to that particular point.

The plumb line does not point towards the center of the Earth as the bob experiences a centrifugal force in addition to the gravitational force due to the Earth. The string must then be adjusted accordingly to balance these forces. Denote $\boldsymbol{g}_0$ as the actual gravitational field ($g_0 \approx 9.81\text{ms}^{-2}$) at the location of the experiment due to the Earth. $\boldsymbol{r}$ is the position vector of the location of the experiment, relative to the origin defined at the center of Earth. The effective gravitational field strength $\boldsymbol{g}$ arises from the vector sum of $\boldsymbol{g}_0$ and the centrifugal force $\boldsymbol{F}_{cen}$ (whose direction is depicted below).

$$\boldsymbol{g} = -g_0\hat{\boldsymbol{r}} + \boldsymbol{F}_{cen} = -g_0\hat{\boldsymbol{r}} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}).$$

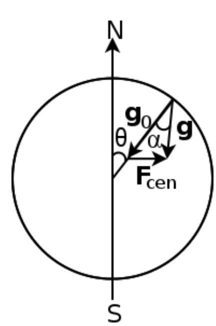


Figure 11.4: Effective gravity at colatitude θ

The radial and tangential components of $\boldsymbol{g}$ are

$$g_r = -(g_0 - F_{cen} \cdot \sin\theta) = -(g_0 - mr\sin^2\theta\omega^2),$$
$$g_t = F_{cen} \cdot \cos\theta = mr\omega^2 \sin\theta\cos\theta.$$

Therefore,

$$\alpha = \tan^{-1}\left|\frac{g_t}{g_r}\right| = \tan^{-1}\frac{r\omega^2\sin\theta\cos\theta}{g_0 - r\omega^2\sin^2\theta}.$$

Incidentally, this means that we have to revamp our definition of the vertical. Usually, the vertical is defined as the direction along $\boldsymbol{g}$ rather than $\boldsymbol{g}_0$ as the latter is difficult to measure (excluding the special case at the equator). Another consequence of the above is that one weighs less near the equator, where g_r is minimum, than at the poles where g_r is maximum.

11.3.1 *Centrifugal Potential*

We can in fact ascribe a potential energy function to the centrifugal force, owing to its conservative nature. Consider the work done by the centrifugal force along a certain path while letting $\boldsymbol{r}$ denote the position vector of a point along the path relative to the origin of the non-inertial frame.

$$\begin{aligned}
W_{cen} &= \int \boldsymbol{F}_{cen} \cdot d\boldsymbol{r} \\
&= \int m\left[(\boldsymbol{\omega}\cdot\boldsymbol{\omega})\boldsymbol{r} - (\boldsymbol{\omega}\cdot\boldsymbol{r})\boldsymbol{\omega}\right]\cdot d\boldsymbol{r} \\
&= \int m\omega^2\boldsymbol{r}\cdot d\boldsymbol{r} - \int m(\boldsymbol{\omega}\cdot\boldsymbol{r})\boldsymbol{\omega}\cdot d\boldsymbol{r} \\
&= \int \frac{1}{2}m\omega^2 d(\boldsymbol{r}\cdot\boldsymbol{r}) - \int \frac{1}{2}md\left[(\boldsymbol{\omega}\cdot\boldsymbol{r})\cdot(\boldsymbol{\omega}\cdot\boldsymbol{r})\right] \\
&= \Delta\left(\frac{1}{2}m\omega^2r^2 - \frac{1}{2}m\left|\boldsymbol{\omega}\cdot\boldsymbol{r}\right|^2\right) \\
&= \Delta\left(\frac{1}{2}m\omega^2r^2\sin^2\theta\right),
\end{aligned}$$

where θ is the angle subtended by $\boldsymbol{\omega}$ and $\boldsymbol{r}$, while r is the distance from the origin of the non-inertial frame. Note that the integral is performed along the entire path that we have chosen. Observe that the work done is path-independent as it is only dependent on the initial and final values of r and θ. Since the change in potential energy between two states is defined as the negative work done by a conservative force along a path between them, the centrifugal potential energy can be defined as

$$U_{cen} = -\frac{1}{2}m\omega^2 r^2 \sin^2\theta, \tag{11.20}$$

by setting the zero reference point at $\theta = 0$. Another way to express the above is $U_{cen} = -\frac{1}{2}m\omega^2 r_\perp^2$, where $r_\perp$ is the perpendicular distance of the point of concern to the axis of rotation. The negative value of U_{cen}, which is amplified with increasing $r_\perp$, implies that particles tend to be propeled away from the axis of rotation in an attempt to minimize their potential energies.

Equatorial Bulge of Earth

The rotation of the Earth causes it to deviate from a perfectly spherical shape as the centrifugal force engenders the stretching of the equator. The equatorial bulge of Earth refers to the difference in the equatorial and polar diameters of Earth and is empirically measured to be 42.7km. A theoretical estimate of this value can be obtained by requiring the surface of the Earth to be equipotential — for the same reasons underscored in the section on tides. The two components of the potential are the gravitational potential due to the Earth and the centrifugal potential. Let R be the radius of the Earth if it were to be perfectly spherical (i.e. we turn off the centrifugal force) and (r, θ) be the polar coordinates of a point on the actual surface of the Earth. Define $h = r - R$ as the excess altitude beyond R — it is presumed that $h \ll R$. In approximating the gravitational potential, the gravitational field strength can be taken to be that at the surface of Earth, g_0, as $h \ll R$, and the zero reference point can be set at $r = R$. The potential at (r, θ) is then

$$g_0 h - \frac{1}{2}\omega^2 r^2 \sin^2\theta = c,$$

where ω is the angular speed of the Earth's rotation about its own axis and c is some constant. Rearranging,

$$h = \frac{c}{g_0} + \frac{\omega^2 r^2 \sin^2\theta}{2g_0}.$$

We can replace r with R here without much penalty as $\frac{\omega^2}{g_0}$ is already small in the case of the Earth.

$$h = \frac{c}{g_0} + \frac{\omega^2 R^2 \sin^2\theta}{2g_0} = b - \frac{\omega^2 R^2 \cos^2\theta}{2g_0},$$

where $b = \frac{c}{g_0} + \frac{\omega^2 R^2}{2g_0}$ is another constant. To set the value of b, we can enforce the condition that the volume of the deformed sphere should still be equal to the volume of the original sphere. This is equivalent to saying that the integral of h over the surface of the original sphere is zero.

$$\int_0^{\pi} \int_0^{2\pi} hR^2 \sin\theta d\phi d\theta = \int_0^{\pi} h 2\pi R^2 \sin\theta d\theta = 0.$$

Substituting the expression for h,

$$\int_0^{\pi} \left(b - \frac{\omega^2 R^2 \cos^2\theta}{2g_0} \right) 2\pi R^2 \sin\theta d\theta = \left[\left(-b\cos\theta + \frac{\omega^2 R^2 \cos^3\theta}{6g_0} \right) 2\pi R^2 \right]_0^{\pi}$$
$$= 2\pi R^2 \left(2b - \frac{\omega^2 R^2}{3g_0} \right).$$

For this integral to be zero,

$$b = \frac{\omega^2 R^2}{6g_0}$$
$$\implies h(\theta) = \frac{\omega^2 R^2}{6g_0} - \frac{\omega^2 R^2 \cos^2\theta}{2g_0} = \frac{\omega^2 R^2}{6g_0} \left(1 - 3\cos^2\theta \right).$$

The excess polar altitude is $h(0) = -\frac{\omega^2 R^2}{3g_0}$ while the excess equatorial altitude is $h(\frac{\pi}{2}) = \frac{\omega^2 R^2}{6g_0}$. The equatorial bulge is thus

$$2h\left(\frac{\pi}{2}\right) - 2h(0) = \frac{\omega^2 R^2}{g_0}.$$

Substituting the actual values ($\omega = 7.29 \times 10^{-5}\text{s}^{-1}$, $R = 6.37 \times 10^6\text{m}$ and $g_0 = 9.81\text{ms}^{-2}$), the hypothetical equatorial bulge is roughly 22.0km, which is the same order of magnitude as the observed value (42.7km) but is still significantly off. The reason behind this non-negligible discrepancy is that in writing the potential as $g_0 h$, we have implicitly assumed that the Earth was spherically symmetric with radius R. However, the whole point of this exercise is to show that it is not! To rectify this issue, one can perform one more iteration to compute a more accurate gravitational potential by treating the actual Earth as a superposition of a uniform sphere of radius R and a shell of height $h(\theta)$ above the sphere and adding the contribution of the latter to $g_0 h$. It turns out that this method would produce a correction factor of roughly $\frac{5}{2}$ which narrows the gap between the predicted and observed equatorial bulges. The remaining discrepancy mainly stems from the fact that the mass of the Earth is not uniformly distributed.

11.3.2 *Rigid Bodies*

An interesting question is to determine the net translational effect of the centrifugal force exerted on an extended body of total mass m. Each point on the body experiences a different centrifugal force so we have to sum up the contributions from all points to compute the net force.

$$\boldsymbol{F}_{cen} = \int [-(\boldsymbol{\omega}\cdot\boldsymbol{r})\boldsymbol{\omega} + (\boldsymbol{\omega}\cdot\boldsymbol{\omega})\boldsymbol{r}]\, dm = -\boldsymbol{\omega}\left(\boldsymbol{\omega}\cdot\int \boldsymbol{r}dm\right) + \omega^2 \int \boldsymbol{r}dm.$$

Since $\int \boldsymbol{r}dm = m\boldsymbol{r}_{CM}$ where $\boldsymbol{r}_{CM}$ is the position vector of the center of mass of the extended body by definition,

$$\boldsymbol{F}_{cen} = -m\boldsymbol{\omega}(\boldsymbol{\omega}\cdot\boldsymbol{r}_{CM}) + m\boldsymbol{r}_{CM}\omega^2 = -m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r}_{CM}).$$

That is, the net centrifugal force on an extended body of total mass m is effectively that on a point mass m located at the center of mass. However, there is no way to ascribe this net centrifugal force to a single point of action — one has to manually integrate the torque experienced by each infinitesimal element due to the centrifugal force to compute the total torque experienced by the body. Despite this limitation, the expression for the net centrifugal force, coupled with $\sum \boldsymbol{F} = m\boldsymbol{a}_{CM}$ for a rigid body, can yield illuminating results in situations where only the translational motion of a rigid body is of concern.

Problem: A uniform thin rod of mass m and length l is constrained to slide within a tube of negligible mass that is rotating at initial angular velocity $\boldsymbol{\omega}_0$ around its center. The center of the rod is initially aligned with the center of the tube. Show that the center of the tube corresponds to an unstable equilibrium for the rod (when their centers coincide) in the direction along the tube. Furthermore, show that the rod will travel to infinity (assuming that the tube is long enough) if it slightly deviates from the center of the tube with a negligible initial velocity.

Let r be the radial distance of the center of the rod from the center of the tube and ω be the instantaneous angular speed of the rod and tube. In the frame rotating at ω with respect to the lab frame (i.e. the tube), the centrifugal force is $mr\omega^2$ outwards. Combining this with the fact that $\sum \boldsymbol{F} = m\boldsymbol{a}_{CM}$ for a rigid body,

$$m\ddot{r} = mr\omega^2$$
$$\ddot{r} = r\omega^2.$$

Therefore, $r = 0$ corresponds to an equilibrium position, regardless of the value of ω. Furthermore, it is an unstable equilibrium as $\ddot{r}$ always has the same sign as r (for all instantaneous ω) such that the rod tends to deviate further away from the center of the tube. In order to solve the above differential equation, we first have to express ω in terms of r via the conservation of angular momentum. When the rod is at radial position r and the tube is rotating at angular velocity ω, the total angular momentum is

$$L = mrv_{CM} + I_{CM}\omega,$$

where $v_{CM} = r\omega$ is the tangential velocity of the center of mass of the rod and $I_{CM} = \frac{1}{12}ml^2$ is the moment of inertia of a uniform rod about its center. Since the initial value of L is $\frac{1}{12}ml^2\omega_0$,

$$\begin{aligned} m\left(r^2 + \frac{l^2}{12}\right)\omega &= \frac{1}{12}ml^2\omega_0 \\ \omega &= \frac{l^2\omega_0}{12r^2 + l^2} \\ \implies \ddot{r} &= \frac{l^4\omega_0^2 r}{(12r^2 + l^2)^2}. \end{aligned}$$

Using the trick $\ddot{r} = \frac{d\dot{r}^2}{2dr}$,

$$\begin{aligned} \int_0^{\dot{r}^2} d(\dot{r}^2) &= \int_0^r \frac{2l^4\omega_0^2 r}{(12r^2 + l^2)^2} dr \\ \dot{r}^2 &= \frac{l^2\omega_0^2}{12} - \frac{l^4\omega_0^2}{12(12r^2 + l^2)} = \frac{l^2\omega_0^2 r^2}{12r^2 + l^2}. \end{aligned}$$

Since $\dot{r}$ must always be positive (in light of $\ddot{r} = r\omega^2$),

$$\dot{r} = \frac{l\omega_0}{\sqrt{12 + \frac{l^2}{r^2}}}.$$

We can directly argue from the above expression that r should tend to infinity. As r increases, $\dot{r}$ increases — thereby leading to a self-perpetuating vicious cycle. The rod thus slides to infinity as an increasing series cannot converge (r is the integral of $\dot{r}$ over time, which is tantamount to summing up individual values of $\dot{r}$ multiplied by small time intervals). This is best illustrated when $r \gg l$ as $\dot{r}$ is approximately constant at $\frac{l\omega_0}{2\sqrt{3}}$.

11.4 Coriolis Force

The Coriolis force is

$$\boldsymbol{F}_{cor} = 2m\boldsymbol{v}_{rot} \times \boldsymbol{\omega}. \tag{11.21}$$

To understand the cause of the Coriolis force intuitively, first consider a person traveling along a straight radial line at a constant outwards velocity v on a merry-go-round in the lab (inertial) frame. The merry-go-round is rotating at angular speed ω anti-clockwise in the lab frame. Since the person is traveling at a constant velocity in the inertial frame, he must not experience any net real force such that his acceleration in the frame of the merry-go-round, if any, must be solely due to the fictitious forces. Suppose the person travels a radial distance $dr = vdt$ from r to $r + dr$ in time dt. Observe that a spot on the carousel at radius r would have traveled distance $r\omega dt$ in that time while a spot on the carousel at radius $r + dr$ would have traveled distance $(r + dr)\omega dt$. Therefore, when the person traverses distance dr, he does not land on the same radial line on the carousel (though he does with respect to a stationary line in the lab frame) — he lies on another one that is displaced by a tangential distance $-dr\omega dt$ (negative as this is clockwise) with respect to the original one. Therefore, there must have been a clockwise tangential acceleration a in the frame of the carousel, which is essentially constant over the small time interval dt. From basic kinematics,

$$\frac{1}{2}a(dt)^2 = -dr\omega dt = -v\omega(dt)^2$$
$$a = -2v\omega.$$

This component of the Coriolis force then accounts for the tangential acceleration in this case (note that v is also the radial velocity in the carousel's frame). Now, a puzzling question to ponder is why the tangential speed only changes by $-v\omega dt$ from $-r\omega$ to $-(r + dr)\omega$ in time dt in the frame of the carousel even though the tangential acceleration is $-2v\omega$. This discrepancy is due to the fact that the radial vector of the person is constantly changing — causing the tangential unit vector to follow suit (this is similar to a common misconception in polar coordinates). We cannot simply integrate a to obtain the tangential velocity as the direction of the tangential acceleration is changing. Instead, we use the following relationship of the tangential acceleration in polar coordinates:

$$a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta},$$

where θ is the angle between the horizontal axis of the rotating frame and the position of the particle. Currently $\dot{\theta} = -\omega$, as the person who does not have an angular velocity in the lab frame possesses a clockwise angular speed ω in the rotating frame of the carousel. Substituting $a_\theta = -2v\omega$, $\dot{r} = v$ and $\dot{\theta} = -\omega$ at the current instance,

$$r\ddot{\theta} = 0,$$

which means that $\dot{\theta}$ remains at $-\omega$ at the next instance. Therefore, the tangential velocity at the next instance is just $-(r + dr)\omega$. But wait, why can we then presume that the distance covered by the person in time dt in the carousel's frame was $\frac{1}{2}a(dt)^2$? Well, this is because a, by proposition, is always tangential to the position vector of the person such that $\frac{1}{2}a(dt)^2$ really represents the distance along an arc (which we equated to $-dr\omega dt$) between two corresponding points at the same radial distance r on two adjacent radial lines of the carousel, that are separated by an angle ωdt. When we compute the increase in tangential velocity, we take the straight line distance between those two points (which is the additional **displacement**[2] covered in time dt due to the increase in tangential velocity) and divide it by the infinitesimal time interval dt. As $dt \to 0$, this straight line distance becomes the length of the arc so that the change in tangential velocity is $\frac{1}{2}adt = -v\omega dt = -dr\omega$, which is consistent with the change from $-r\omega$ to $-(r + dr)\omega$.

Now, let us consider a different set-up to elucidate the radial component of the Coriolis force. The person now travels at an instantaneous tangential speed v anti-clockwise at distance r in the frame of the carousel — symbolizing that the person travels at an anti-clockwise tangential speed $v+r\omega$ in the lab frame. It is presumed that the person is again free from real forces. In the lab frame, after time dt, the radial coordinate would have changed from r to

$$\begin{aligned} r + dr &= \sqrt{r^2 + (v + r\omega)^2(dt)^2} = r\sqrt{1 + \frac{(v + r\omega)^2}{r^2}(dt)^2} \\ &\approx r\left(1 + \frac{(v + r\omega)^2}{2r^2}(dt)^2\right), \\ dr &= \frac{1}{2}\left(\frac{v^2}{r} + 2v\omega + r\omega^2\right)(dt)^2. \end{aligned}$$

Then, from basic kinematics again (with $\dot{r} = 0$ initially and with $\ddot{r}$ approximately constant over dt),

$$\begin{aligned} \frac{1}{2}\ddot{r}(dt)^2 &= \frac{1}{2}\left(\frac{v^2}{r} + 2v\omega + r\omega^2\right)(dt)^2 \\ \implies \ddot{r} &= \frac{v^2}{r} + 2v\omega + r\omega^2. \end{aligned}$$

[2] Note that this is not the distance.

Observe that $\ddot{r}$ must be the same in both frames as their origins are relatively stationary. Now, let us consider the perspective of the carousel. The radial forces here are the centrifugal force $mr\omega^2$ and possibly the Coriolis force. Applying $F = ma$ in polar coordinates,

$$mr\omega^2 + F_{cor} = m\ddot{r} - m\frac{v^2}{r},$$

where $\frac{v^2}{r}$ is the instantaneous centripetal acceleration term (of the person) in the carousel's frame (do not confuse this with $r\omega^2$). Equating the two above expressions for $\ddot{r}$, we find that there must be a radial Coriolis force

$$F_{cor} = 2mv\omega.$$

Now that we have obtained an intuitive feeling for the advent of the Coriolis force, there is an interesting analogy between it and the magnetic force experienced by a charge. $2m$ corresponds to the charge while $\boldsymbol{\omega}$ is akin to the magnetic field. This analogy is not profound in any sense but it helps us to determine the direction of the Coriolis force. It is always perpendicular to the instantaneous velocity of a particle in a rotating frame and hence does no work. Furthermore, if $\boldsymbol{\omega}$ is pointing out of the page, all particles are deflected towards the right (in the page) of their instantaneous velocities, regardless of their exact positions and velocities. Otherwise if $\boldsymbol{\omega}$ points into the page, all particles are deflected towards the left of their instantaneous velocities. An intriguing application of this analogy is illustrated below.

Problem: A $-k\boldsymbol{r}$ force is exerted on a particle of mass m where $\boldsymbol{r}$ is the position vector of the particle from the origin. Show that under arbitrary initial conditions, the trajectory of the particle is generally an ellipse. You do not have to calculate the specific parameters of the ellipse.

The first observation is that the angular momentum of the particle must be conserved as the force is radial — a typical property of any central force problem. Hence, the motion of the particle is restricted to the plane perpendicular to the angular momentum vector — reducing this problem to two dimensions. Instead of solving a particular case of the central force problem head-on, we can switch to a new frame S' that rotates at a certain angular velocity $\boldsymbol{\omega}$ (perpendicular to the plane of motion) with respect to the lab frame S such that the centrifugal force exactly cancels the central $-k\boldsymbol{r}$ force. When $\boldsymbol{r}$ is perpendicular to $\boldsymbol{\omega}$, the centrifugal force becomes $m\omega^2\boldsymbol{r}$ by Eq. (11.19). The particle's equation of motion in S' is thus

$$m\boldsymbol{a}_{rot} = -k\boldsymbol{r} + m\omega^2\boldsymbol{r} + 2m\boldsymbol{v}_{rot} \times \boldsymbol{\omega}.$$

Choosing

$$\omega = \sqrt{\frac{k}{m}}$$

causes the centrifugal force to cancel the central force at all times. Then,

$$m\boldsymbol{a}_{rot} = 2m\boldsymbol{v}_{rot} \times \boldsymbol{\omega}.$$

This equation of motion is analogous to that of a charged particle in a uniform magnetic field and describes circular motion, as the acceleration is always perpendicular to the instantaneous velocity — thus maintaining its magnitude and providing a centripetal acceleration. The angular frequency Ω of this circular motion (refer to the chapter on magnetism) can be computed from the fact that a_{rot} is the centripetal acceleration for a constant speed v_{rot}; $a_{rot} = v_{rot}|\Omega|$.

$$mv_{rot}|\Omega| = 2mv_{rot}|\omega|$$

$$\implies |\Omega| = 2|\omega|.$$

Now, what is the direction of $\boldsymbol{\Omega}$? Suppose that the plane of motion of the particle is this page and $\boldsymbol{\omega}$ points out of the page (anti-clockwise). A particle with any position and velocity is deflected rightwards of their instantaneous velocity — implying that it tends to travel clockwise. Hence, $\boldsymbol{\Omega} = -2\boldsymbol{\omega}$. The radius of circular motion R can easily be computed as

$$R = \frac{v_{rot}}{\Omega} = \frac{v_{rot}}{2\omega},$$

where v_{rot} is the constant speed in S' (it can be calculated by $|\boldsymbol{v}_0 + \boldsymbol{\omega} \times \boldsymbol{r}_0|$ where $\boldsymbol{v}_0$ and $\boldsymbol{r}_0$ are the initial velocity in S and the initial position vector of the particle respectively). Since the particle undergoes circular motion at $\boldsymbol{\Omega}$ about a certain center C in S', its coordinates (x', y') obey

$$x' + iy' = C_0 + Re^{i(\Omega t - \phi_1)},$$

where C_0 is a certain complex number. To see why this is so, simply imagine this complex number in the complex plane. The real part represents the x' coordinate while the complex part corresponds to the y' coordinate. Evidently, the trajectory of the complex number is a circle of radius R about the center at coordinate C_0 (with angular velocity Ω) and corresponds to the trajectory of the particle. ϕ_1 delineates the initial position of the particle with respect to the center. For the sake of convenience, express C_0 in polar form

$$C_0 = ce^{i\phi_2},$$

where c is the distance of the center of rotation from the origin in S' (which overlaps with the origin in S). The computation of c is rather straightforward, albeit tedious, and is left to the reader.

$$x' + iy' = ce^{i\phi_2} + Re^{i(\Omega t - \phi_1)}.$$

To transform these coordinates in S' into those in S, observe that the axes of S' would have rotated an anti-clockwise angle ωt with respect to those of S at time t. Thus, a vector that makes an anti-clockwise angle θ with respect to the x'-axes at time t would make angle $\theta + \omega t$ with respect to the x-axes. We simply have to multiply $x' + iy'$ by $e^{i\omega t}$ to obtain the complex representation $x + iy$ in S, where (x, y) are the instantaneous coordinates of the particle in S.

$$x + iy = (x' + iy')e^{i\omega t} = \left(ce^{i\phi_2} + Re^{i(\Omega t - \phi_1)}\right) e^{i\omega t} = ce^{i(\omega t + \phi_2)} + Re^{-i(\omega t + \phi_1)}$$

as $\Omega = -2\omega$. The x and y coordinates in S are the real and complex components of this complex number, respectively.

$$x = c\cos(\omega t + \phi_2) + R\cos(\omega t + \phi_1),$$

$$y = c\sin(\omega t + \phi_2) - R\sin(\omega t + \phi_1).$$

It is hard to see why these parametric equations represent an ellipse at the first glance so consider the special case where $\phi_1 = \phi_2 = \phi$ for now. Then,

$$x = (R + c)\cos(\omega t + \phi),$$

$$y = (c - R)\sin(\omega t + \phi)$$

$$\implies (R - c)^2 x^2 + (R + c)^2 y^2 = (R^2 - c^2)^2.$$

Assuming $R \neq c$ (one can show that this occurs only if the initial velocity in S is purely radial),

$$\frac{x^2}{(R + c)^2} + \frac{y^2}{(R - c)^2} = 1,$$

which is the standard equation of an ellipse. Now, what about the case where $\phi_1 \neq \phi_2$? Consider the multiplication of $x + iy$ with $e^{-i(\frac{\phi_1 - \phi_2}{2})}$ which represents the transformation of coordinates from the axes in S to those with respect to a new set of axes in S" that are rotated $\frac{\phi_1-\phi_2}{2}$ anti-clockwise with respect to the former.

$$(x+iy)e^{-i(\frac{\phi_1-\phi_2}{2})} = ce^{i(\omega t+\frac{\phi_1+\phi_2}{2})} + Re^{-i(\omega t+\frac{\phi_1+\phi_2}{2})}.$$

Since the angles in the exponents are equal, the trajectory of the particle with respect to the new set of axes in S" is an ellipse, as implied by the previous result. Therefore, the trajectory of the particle in S is a tilted ellipse with semi-axes lengths $R+c$ and $|R-c|$, rotated by an anti-clockwise angle $\frac{\phi_1-\phi_2}{2}$ with respect to the positive x-axis.

11.4.1 *Deflection of Free-Falling Objects*

A prominent manifestation of the Coriolis effect is the eastwards deflection of a projectile near the surface of Earth. Let $\boldsymbol{\omega}$ be the angular velocity of Earth and $\boldsymbol{r}$ denote the instantaneous vector pointing from the center of Earth to the projectile — we assume that the range of the projectile is small relative to the radius of Earth such that its colatitude and $\hat{\boldsymbol{r}}$ remain constant. The projectile's equation of motion in Earth's rotating frame is

$$m\boldsymbol{a} = -mg_0\hat{\boldsymbol{r}} - m\boldsymbol{\omega}\times(\boldsymbol{\omega}\times\boldsymbol{r}) + 2m\boldsymbol{v}\times\boldsymbol{\omega},$$

where we have dropped the subscript "rot". In this analysis, we will be ignoring terms that are second order in ω so the centrifugal term is negligible. Then,

$$m\boldsymbol{a} = -mg_0\hat{\boldsymbol{r}} + 2m\boldsymbol{v}\times\boldsymbol{\omega}.$$

Integrating the above,

$$\boldsymbol{v} = \boldsymbol{v}_0 - g_0t\hat{\boldsymbol{r}} + 2(\boldsymbol{r}-\boldsymbol{r}_0)\times\boldsymbol{\omega},$$

where $\boldsymbol{v}_0$ and $\boldsymbol{r}_0$ are the initial velocity and position of the object. Substituting this expression for $\boldsymbol{v}$ into $\boldsymbol{a}$,

$$\boldsymbol{a} = -g_0\hat{\boldsymbol{r}} + 2\left[\boldsymbol{v}_0 - g_0t\hat{\boldsymbol{r}} + 2(\boldsymbol{r}-\boldsymbol{r}_0)\times\boldsymbol{\omega}\right]\times\boldsymbol{\omega}.$$

Neglecting the second-order term in ω,

$$\boldsymbol{a} = -g_0\hat{\boldsymbol{r}} + 2\boldsymbol{v}_0\times\boldsymbol{\omega} - 2g_0t\hat{\boldsymbol{r}}\times\boldsymbol{\omega}.$$

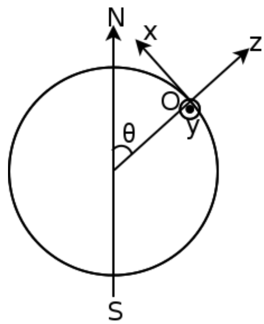


Figure 11.5: Coordinate system at colatitude θ

Integrating this again and using the fact that $\boldsymbol{v} = \boldsymbol{v}_0$ initially,

$$\boldsymbol{v} = \boldsymbol{v}_0 - g_0 t\hat{\boldsymbol{r}} + 2t\boldsymbol{v}_0 \times \boldsymbol{\omega} - g_0 t^2 \hat{\boldsymbol{r}} \times \boldsymbol{\omega}.$$

Integrating this one last time and using the initial condition $\boldsymbol{r} = \boldsymbol{r}_0$ at $t = 0$,

$$\boldsymbol{r} = \boldsymbol{r}_0 + \boldsymbol{v}_0 t - \frac{1}{2} g_0 t^2 \hat{\boldsymbol{r}} + t^2 \boldsymbol{v}_0 \times \boldsymbol{\omega} - \frac{1}{3} g_0 t^3 \hat{\boldsymbol{r}} \times \boldsymbol{\omega}. \tag{11.22}$$

To visualize this result, consider the plane containing $\boldsymbol{\omega}$ and $\boldsymbol{r}_0$ with $\boldsymbol{\omega}$ being the vertical axis, pointing from the South pole to the North pole (Fig. 11.5). Define the origin at the surface of the Earth at which $\boldsymbol{r}_0$ intersects, the positive z-axis to be aligned with $\hat{\boldsymbol{r}}_0$ (such that $\hat{\boldsymbol{r}} = \hat{\boldsymbol{k}}$) and the positive x-axis to point tangentially northwards in this plane. Then, the positive y-axis can be defined from the fact that $\hat{\boldsymbol{i}} \times \hat{\boldsymbol{j}} = \hat{\boldsymbol{k}}$ for a conventional coordinate system and is directed westwards (opposite to the direction that the origin tends to rotate towards, and is out of the page in this case). This coordinate system is fixed to the Earth (i.e. rotating).

Besides the first three terms, $\boldsymbol{r}_0 + \boldsymbol{v}_0 t - \frac{1}{2} g_0 t^2 \hat{\boldsymbol{r}}$, that we expect of a standard projectile motion, there is an additional expression $t^2 \boldsymbol{v}_0 \times \boldsymbol{\omega} - \frac{1}{3} g_0 t^3 \hat{\boldsymbol{r}} \times \boldsymbol{\omega}$. The latter $-\frac{1}{3} g_0 t^3 \hat{\boldsymbol{r}} \times \boldsymbol{\omega}$ term is dominant in the long run and yields an additional $\frac{1}{3} g_0 t^3 \omega \sin\theta$ displacement eastwards (into the page) where θ is the approximately constant angle between $\boldsymbol{\omega}$ and $\boldsymbol{r}$.

11.4.2 *Foucault Pendulum*

The Foucault pendulum is a famous device that demonstrates the rotation of Earth. Due to the rotation of the Earth, the plane of oscillation of a swinging pendulum precesses in the frame of the Earth, at a rate that is dependent on the colatitude θ at which the experiment is conducted. This

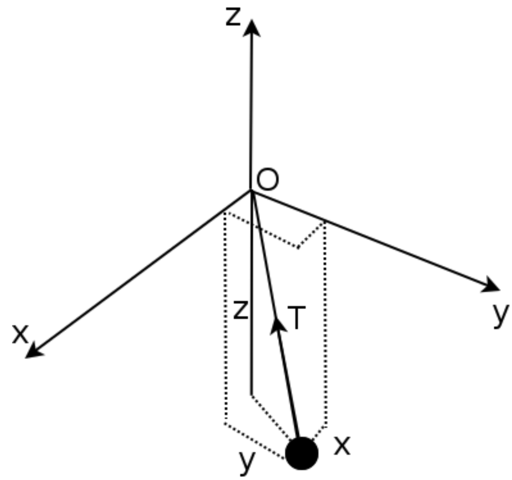


Figure 11.6: Coordinate system with pendulum

is obvious when $\theta = 0$ as an external inertial observer (possibly located on the distant, fixed stars) would observe the pendulum to swing in a single plane while Earth rotates at angular speed ω beneath it. Thus, in the frame of the Earth, the plane of oscillation of the pendulum should precess at $-\omega$.

Moving on to the general case while adopting the coordinate system defined in Fig. 11.5, $\boldsymbol{\omega} = (\omega \sin\theta, 0, \omega\cos\theta)$.

In the frame of the Earth, define the origin at the fixed point of the rope holding the pendulum. Let the coordinates of the pendulum be (x, y, z). Then, if T is the tension in the string, its components are $(-\frac{Tx}{l}, -\frac{Ty}{l}, -\frac{Tz}{l})$ where l is the length of the string (Fig. 11.6). For small oscillations, the rope does not deviate much from the z-axis such that x and y are small and z is approximately $-l$. The tension vector is then approximately $\boldsymbol{T} = (-\frac{Tx}{l}, -\frac{Ty}{l}, T)$. The equation of motion of the pendulum in the Earth's frame is

$$m\boldsymbol{a} = -mg_0\hat{\boldsymbol{k}} + 2m\boldsymbol{v} \times \boldsymbol{\omega} + \boldsymbol{T},$$

where we have neglected the centrifugal term. For $z \approx -l$, $\dot{z}$ and $\ddot{z}$ are both small. This, coupled with the fact that the z-component of the Coriolis term disappears when $\dot{y}$ is small, shows that the z-component of $\boldsymbol{T}$ must nullify $-mg_0\hat{\boldsymbol{k}}$ so that $\ddot{z} = 0$. Then,

$$T = mg_0.$$

The equation of motion becomes

$$\boldsymbol{a} = \begin{pmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{pmatrix} = \begin{pmatrix} -\frac{g_0 x}{l} \\ -\frac{g_0 y}{l} \\ 0 \end{pmatrix} + 2 \begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} \times \begin{pmatrix} \omega \sin\theta \\ 0 \\ \omega \cos\theta \end{pmatrix}.$$

In terms of its components,

$$\begin{aligned} \ddot{x} &= -\frac{g_0}{l} x + 2\omega \cos\theta \dot{y}, \\ \ddot{y} &= -\frac{g_0}{l} y - 2\omega \cos\theta \dot{x}, \\ \ddot{z} &\approx 0, \end{aligned}$$

for small $\dot{y}$ and $\dot{z}$. Now, we define a new complex variable $\eta = x + iy$ and add the second equation multiplied by i to the first. Defining $\omega_0^2 = \frac{g_0}{l}$ as the natural frequency of the pendulum,

$$\ddot{\eta} = -\omega_0^2 \eta - 2i\omega \cos\theta \dot{\eta}.$$

Guessing an exponential solution of the form $\eta = \eta_0 e^{\alpha t}$ and solving for α,

$$\alpha^2 = -\omega_0^2 - 2i\omega \cos\theta \alpha,$$

whose solutions are

$$\alpha = -i\omega \cos\theta \pm \sqrt{-\omega^2 \cos^2\theta - \omega_0^2}.$$

In practice, $\omega_0 \gg \omega$ (as the rotation of Earth about its own axis is slow) such that

$$\alpha \approx -i\omega \cos\theta \pm i\omega_0.$$

The general solution for η is thus

$$\eta = Ae^{(-i\omega\cos\theta + i\omega_0)t} + Be^{-(i\omega\cos\theta + i\omega_0)t} = e^{-i\omega\cos\theta t}\left(Ae^{i\omega_0 t} + Be^{-i\omega_0 t}\right)$$

for some complex constant A and B. Rewriting the above in terms of sine and cosine,

$$\eta = e^{-i\omega\cos\theta t}(C \sin\omega_0 t + D \cos\omega_0 t),$$

where C and D are new complex constants. Recall that η on an Argand diagram represents the trajectory of the pendulum (x is the real component while y is the imaginary component). Since $\omega_0 \gg \omega$, the term in front of the brackets is essentially constant for a small period of time. Then, η is solely described by the rapidly oscillating terms in $\omega_0 t$ — implying that the x and y components of the pendulum vary sinusoidally at the natural frequency ω_0. After a significant amount of time Δt, the $e^{-i\omega\cos\theta t}$ term comes into effect and rotates the plane of oscillation by $-\omega\cos\theta\Delta t$ anti-clockwise. Thus, the pendulum oscillates at angular frequency ω_0 while its plane of oscillation slowly precesses at angular frequency

$$\Omega = -\omega\cos\theta.$$

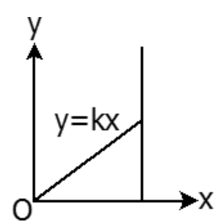
Problems

1. *Pendulum**

Supposing that you are trapped in an unknown room at the North pole with a constant acceleration, how would you determine the acceleration of the room using a pendulum, a protractor and a stopwatch? Suppose that you know that the true gravity points in the direction perpendicular to the ground of the room.

2. *Cup of Water***

Supposing that you are trapped in an unknown room at the North pole with a constant acceleration, you observe that the surface of a cylindrical cup of water (that is stationary relative to you) obeys the equation $y = kx$ from the side-view where y is the height of the liquid level from the base, x is the horizontal distance from the left tip of the base and k is a constant. The length of the base is l. Suppose that you know that the true gravity points in the direction perpendicular to the ground of the room. Explain how you can determine your acceleration using this cup, a knife and a stopwatch (it is also possible to do so with a cup and a ruler, albeit involving much more tedious calculations). Hint: perform another experiment with the cup.



3. *Rolling without Slipping**

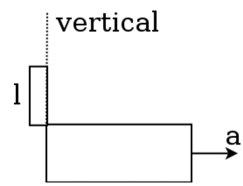
A uniform sphere of radius R lies on a rough plank. If the plank is given a uniform acceleration A rightwards and the sphere subsequently rolls without slipping, what is the angular acceleration of the sphere?

4. *Falling Off an Accelerating Circle**

A particle initially rests on top of a stationary, massive circle. If the circle is now accelerated at a constant acceleration a leftwards, determine the angle θ from the vertical at which the particle loses contact with the circle.

5. *Swinging Door*

The door of a truck, which can be modeled as a uniform rod of mass m and length l, is initially fully open, as shown in the diagram below. The truck then experiences a constant acceleration a rightwards. Determine the angular velocity of the door $\dot{\theta}$, as a function of its angle θ from the vertical in the diagram. Finally, determine the force exerted by the hinge on the door as a function of θ.



Centrifugal and Coriolis Forces

6. *Rotating Candle**

A candle is placed inside a transparent box. The box is then rotated by an external agency about a fixed center of rotation on a horizontal table at a constant angular velocity ω. What direction does the flame tilt towards after a long time (assuming that there is sufficient oxygen supply)?

7. *Circular Motion in Rotating Frame**

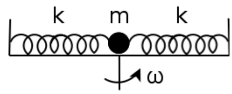
A carousel is rotating at angular velocity ω_1 anti-clockwise in the lab frame about its center. In a frame fixed to the carousel, a particle of mass m undergoes circular motion at radius R and angular velocity ω_2 anti-clockwise, also about the center of the carousel. Explain why the friction force on the particle should be $mR(\omega_1+\omega_2)^2$ from both the perspectives of the lab frame and the rotating frame.

8. *Rotating Bucket**

Show, without the consideration of any forces, that the surface of a cylinder of water rotating about its cylindrical axis at an angular velocity ω is a paraboloid at equilibrium.

9. *Mass on Fork**

A mass m is placed at the center of a frictionless fork as shown in the figure on the next page. It is connected to two identical massless springs of spring constant k. If the fork is now rotated anti-clockwise at angular velocity ω about a vertical axis passing through the center of the fork, determine the range of ω for which m can undergo simple harmonic motion when it is slightly displaced along the axial direction of the spring. Also determine the angular frequency of such oscillations.



10. *Disk on Rotating Table**

A large horizontal table is rotating at a constant anti-clockwise angular velocity ω relative to the upwards direction, about a vertical axis crossing a certain point O on the table. A uniform disk is then placed on the table and initially possesses arbitrary center of mass and angular velocities. If the surface between the table and the disk is rough, determine the steady state angular velocity of the disk. Assume that the normal force on the disk is evenly distributed over the disk.

11. *Straight Line on Carousel**

A frictionless carousel is rotating at a constant angular velocity ω anti-clockwise in the lab frame. Consider a non-rotating coordinate system x and y along the plane of the carousel in the lab frame. A particle is placed at an initial coordinate (x_0, y_0) and given an initial velocity (u_x, u_y) along the plane of the carousel. In the lab frame, the trajectory of the particle is a straight line. Determine the coordinates of the particle in a frame fixed to the carousel as a function of time by working in the rotating frame. Show that the trajectory of the particle in this frame, after a significant amount of time, takes the form of a spiral. The simple way is to directly transform the coordinates of the particle in the lab frame into those of the rotating frame, but try to solve this problem via the equation of motion in the rotating frame.

12. *Larmor's Theorem**

An isolated system of charges is currently interacting in the lab frame. Show that if a small magnetic field is introduced, the system evolves in the same way, except that the entire system now precesses in a certain plane at a certain angular velocity. Hint: consider a rotating frame and try to cancel some terms in the equation of motion of a charge.

13. *Sniping South***

A sniper at colatitude θ wishes to shoot a target along the south direction. Suppose that the bullet leaves his gun barrel at an initial velocity v_0, what is the southwards distance between his target and him, tangential to the surface of the Earth at his location, for which he can simply point his gun towards the south? Neglect the drop of the bullet towards the ground (well, perhaps the target is very tall) and assume that the target is not too far away such that its colatitude is also θ.

Solutions

1. Pendulum*

Firstly, observe the equilibrium position of the pendulum. If the pendulum bob tilts from the vertical, there must be a horizontal component of acceleration. Consider the plane containing the vertical and the pendulum string and define the direction of the horizontal axis to be positive towards the bob. Denote θ as the positive angle between the equilibrium position of the string and the vertical. The forces on the pendulum bob are the fictitious inertial force, tension and gravity. Let $-a_x$ and a_y denote the horizontal and vertical components of the acceleration of the room, where the latter is positive when directed upwards. Then, the bob experiences inertial forces ma_x in the positive horizontal direction and ma_y downwards. In order for the bob to remain at equilibrium, the combination of the inertial force and gravity must be parallel to the tension (i.e. the string), Hence,

$$\tan\theta = \frac{a_x}{a_y + g}.$$

Next, displace the pendulum from the equilibrium position slightly and time the period of small oscillations T. The angular frequency is $\omega = \frac{2\pi}{T}$. Since the pendulum effectively lives in a world with gravitational field strength $\sqrt{a_x^2 + (a_y+g)^2}$, directed θ clockwise from the vertical, the angular frequency of the pendulum should be

$$\omega = \sqrt{\frac{\sqrt{a_x^2 + (a_y+g)^2}}{l}},$$

where l is the length of the string. Substituting the first equation into the second,

$$(g + a_y)\sec\theta = \omega^2 l$$
$$a_y = \omega^2 l\cos\theta - g,$$
$$a_x = (g + a_y)\tan\theta = \omega^2 l\sin\theta.$$

2. Cup of Water**

Consider the plane of the side view of the cup and define the x and y-axes to be positive rightwards and upwards respectively. Let the horizontal and vertical accelerations of the room be $-a_x$ and a_y respectively. Then, every fluid element of mass dm in the cup experiences inertial forces dma_x rightwards and dma_y downwards. The water in the cup effectively lives in a world where the uniform gravitational field is $(a_x, -(g+a_y))$. Now, observe that the effective gravity must be perpendicular to a liquid surface at equilibrium as a liquid surface cannot withstand any shear forces. This requires

$$\frac{g + a_y}{a_x} = \frac{1}{k}$$
$$\implies k(g + a_y) = a_x.$$

Another equivalent way of obtaining this result is to observe that the surface of the water must correspond to an equipotential surface (due to the potential energies ascribed to the inertial force and the gravitational force). The potential energy per unit mass at a point on the surface at coordinates (x, kx) is $-a_x x + k(g + a_y)x$ and must undertake a common value 0 (potential energy per unit mass at the origin) — leading to the same conclusion. Now, let us perform another experiment with the cup. By using the knife to poke a small hole on the surface of the cup and observing the resultant parabolic motion,[3] we can deduce another relationship between a_y and a_x. It is convenient to puncture the midpoint of the fluid level at the right edge (at height $\frac{kl}{2}$). Then, record the time t taken for the stream of water to reach the table that the cup is sitting on. Since the effective vertical gravity is $g + a_y$, basic kinematics yields

$$\frac{kl}{2} = \frac{1}{2}(g + a_y)t^2,$$

as the initial vertical velocity of the water escaping the cup should be zero. Then,

$$a_y = \frac{kl}{t^2} - g,$$
$$a_x = \frac{k^2 l}{t^2}.$$

You can also determine a relationship between a_x and a_y in the second experiment by solely recording the trajectory of the fluid, but the ensuing calculations will be much more tedious.

3. Rolling without Slipping*

Define the x and y-axes to be positive rightwards and upwards. Let the mass of the sphere be m. In the accelerating frame of the plank, the sphere

[3]Note that the symmetrical axis of the parabola is no longer along the vertical as the effective gravity does not point along this direction.

experiences an inertial force $-mA$ (leftwards), friction f, its weight $-mg$ (downwards) and a normal force due to the plank $N = mg$ (upwards to nullify the weight). The acceleration a of its center of mass is

$$f - mA = ma.$$

Furthermore, the friction produces a torque about the center of the sphere which engenders an anti-clockwise angular acceleration α.

$$Rf = \frac{2}{5}mR^2\alpha,$$

where we have used the fact that the moment of inertia of a uniform sphere of mass m and radius R about an axis passing through its center is $\frac{2}{5}mR^2$. For the sphere to not slip with respect to the plank, the acceleration of the bottom of the sphere (point of contact) must be zero in the plank's frame. Therefore,

$$a + R\alpha = 0.$$

Substituting $a = -R\alpha$ into the first equation,

$$mR\alpha = mA - f.$$

Solving this equation simultaneously with the second equation,

$$\alpha = \frac{5A}{7R},$$

which is an expected result, as the situation in the plank's frame is akin to a circle rolling without slipping along an inclined plane with an angle of inclination of 90° (i.e. a vertical wall) and g being A.

4. Falling Off an Accelerating Circle*

Define the x and y-axes to be positive rightwards and upwards and the origin to be at the center of the circle. In the accelerating frame of the circle, the particle experiences an inertial force ma rightwards, where m is its mass. As this inertial force is conservative, we can ascribe an inertial potential energy $-max$ to the particle where x is its x-coordinate. When the position vector of the particle subtends an angle θ with the vertical, its total potential energy is

$$U = mgR\cos\theta - maR\sin\theta,$$

where R is the radius of the circle, as the coordinates of the particle are $(R\sin\theta, R\cos\theta)$. Applying the conservation of energy, the kinetic energy of the particle at angle θ is

$$\frac{1}{2}mR^2\dot{\theta}^2 = mgR(1-\cos\theta) + maR\sin\theta$$
$$\implies mR\dot{\theta}^2 = 2mg(1-\cos\theta) + 2ma\sin\theta,$$

where $\dot{\theta}$ is the angular velocity of the particle. Now, at angle θ, the combination of the gravitational, inertial and normal forces provides the necessary centripetal force, such that

$$mg\cos\theta - N - ma\sin\theta = mR\dot{\theta}^2.$$

When the particle loses contact with the circle, $N = 0$. Substituting the previous expression for $mR\dot{\theta}^2$,

$$mg\cos\theta - ma\sin\theta = 2mg(1-\cos\theta) + 2ma\sin\theta$$
$$3g\cos\theta - 3a\sin\theta = 2g$$
$$3\sqrt{g^2+a^2}\cos\left(\theta + \tan^{-1}\frac{a}{g}\right) = 2g$$
$$\theta = \cos^{-1}\frac{2g}{3\sqrt{g^2+a^2}} - \tan^{-1}\frac{a}{g}.$$

5. Swinging Door**

Define the x and y-axes to be positive rightwards and upwards and the origin to be at the hinge. In the frame of the uniformly accelerating truck, an inertial force acts on the rod. Since the inertial force is akin to the force due to a uniform gravitational field, an effective inertial force $-ma$ acts on the center of mass of the rod along the x-direction. Furthermore, since the inertial force is conservative, we can ascribe to it a potential energy function max where x is the x-coordinate of the center of mass (similar to the gravitational potential). Applying the conservation of energy to the initial state of the door and its state at angle θ,

$$\frac{mal\sin\theta}{2} = \frac{1}{2}\cdot\frac{1}{3}ml^2\dot{\theta}^2,$$

where $\frac{1}{3}ml^2$ is the moment of inertia of a uniform rod of mass m and length l about one of its ends. Simplifying,

$$\dot{\theta}^2 = \frac{3a}{l}\sin\theta$$
$$\dot{\theta} = \sqrt{\frac{3a}{l}\sin\theta}.$$

Now, to determine the force exerted by the hinge, we have to ensure that the net force on the door correctly reflects the acceleration of the center of mass. The latter is $\frac{l}{2}\ddot{\theta}$ tangentially and $-\frac{l}{2}\dot{\theta}^2$ radially (centripetal acceleration). We can compute $\ddot{\theta}$ using the trick

$$\ddot{\theta} = \frac{d\dot{\theta}^2}{2d\theta} = \frac{3a}{2l}\cos\theta.$$

The acceleration of the center of mass is thus

$$\boldsymbol{a}_{CM} = \left(\frac{3a}{2} - \frac{9a}{4}\cos^2\theta\right)\hat{\boldsymbol{\imath}} - \frac{9a}{4}\sin\theta\cos\theta\hat{\boldsymbol{\jmath}}.$$

The forces on the door are the inertial force and that due to the hinge.

$$\boldsymbol{F}_{hinge} - ma\hat{\boldsymbol{\imath}} = m\boldsymbol{a}_{CM}$$
$$\boldsymbol{F}_{hinge} = \frac{(10 - 9\cos^2\theta)ma}{4}\hat{\boldsymbol{\imath}} - \frac{9ma\sin\theta\cos\theta}{4}\hat{\boldsymbol{\jmath}}$$

6. Rotating Candle*

Consider a rotating frame with its origin at the center of rotation that possesses an angular velocity ω with respect to the lab frame. The box is stationary in this frame but its constituents experience a centrifugal force (directed radially outwards). Therefore, there must be a pressure gradient within the box to counteract the centrifugal force on each air section for the air inside to attain equilibrium. Specifically, the pressure in the box must increase radially outwards to produce a net force radially inwards. Therefore, the flame tilts radially inwards — towards a region of lower pressure.

7. Circular Motion in Rotating Frame*

We first begin with the simpler case — the lab frame. Since angular velocities add, the particle simply rotates at angular velocity $\omega_1 + \omega_2$ anti-clockwise in the lab frame. Hence, the friction force must be

$$f = mR(\omega_1 + \omega_2)^2$$

to provide the necessary centripetal force. In the frame rotating at ω_1 anti-clockwise relative to the lab frame, the particle first experiences a centrifugal force $mR\omega_1^2$ radially outwards. Furthermore, since the particle is traveling at velocity $R\omega_2$ anti-clockwise, it experiences a Coriolis force $2mR\omega_1\omega_2$ radially outwards. The combination of friction and the two fictitious forces must provide the necessary centripetal force $mR\omega_2^2$ radially inwards for the particle to undergo circular motion at angular velocity ω_2.

$$f - mR\omega_1^2 - 2mR\omega_1\omega_2 = mR\omega_2^2$$
$$f = mR(\omega_1 + \omega_2)^2.$$

8. Rotating Bucket*

In the rotating frame of the bucket, the fluid surface must be equipotential as the fluid cannot withstand any shear forces (remember to include centrifugal potential energy). Equating the potential energy per unit mass at the surface at a perpendicular distance r from the axis of the bucket to the potential energy per unit mass at the surface along the axis of the bucket,

$$gy - \frac{r^2\omega^2}{2} = gy_0,$$

where y is the height of the surface at perpendicular distance r and y_0 is the equilibrium height along the axis of the bucket. Rearranging,

$$y = y_0 + \frac{r^2\omega^2}{2g}.$$

9. Mass on Fork*

Suppose that m is displaced from the center by a distance x along the axial direction. Letting l_0 be the rest length of the springs and l be the distance between the fixed end of a spring and the center of the fork, the potential energy of m is

$$U = \frac{1}{2}k(l + x - l_0)^2 + \frac{1}{2}k(l - x - l_0)^2 - \frac{mx^2\omega^2}{2},$$

where we have included the centrifugal potential. Computing the derivatives of U with respect to x,

$$\frac{dU}{dx} = k(l + x - l_0) - k(l - x - l_0) - mx\omega^2 = 2kx - mx\omega^2$$

$$\frac{d^2U}{dx^2} = 2k - m\omega^2.$$

If m is indeed able to oscillate about the center of the fork, its angular frequency of oscillations is

$$\Omega = \sqrt{\frac{\frac{d^2U}{dx^2}\big|_{x=0}}{m}} = \sqrt{\frac{2k}{m} - \omega^2}.$$

m will be in a neutral or unstable equilibrium at $x = 0$ if the term in the square root is not positive. This requires

$$\omega \geq \sqrt{\frac{2k}{m}}.$$

Otherwise if $\omega < \sqrt{\frac{2k}{m}}$, m will undergo simple harmonic motion with the above angular frequency Ω.

10. Disk on Rotating Table**

In such problems where a rough platform is moving, it is always intuitive to consider the perspective of the platform as the direction of friction depends on the relative velocity between the platform and an object placed on it. Therefore, we can switch to the frame of the table which is rotating at anti-clockwise angular velocity ω relative to the lab frame. In this frame, the disk experiences centripetal and Coriolis forces in addition to friction.

We shall prove that the former two factors result in zero torque about the center of the disk. Referring to the diagram on the left in Fig. 11.7, join a line from O to the center of the disk O'. Observe that the torques about O' due to the centripetal forces on two points on the disk that are mirror images about this line nullify each other. An example of such a pair is depicted in the same figure on the left. Since the line divides the circle into two equal halves, the centripetal force must contribute to zero net torque about O'. To analyze the effect of the Coriolis force and friction, it is important to note that even though the velocity of the center of the disk may not have attained a steady value when the steady state angular velocity is achieved, the motion of the disk can be represented by a pure rotation about an instantaneous center of

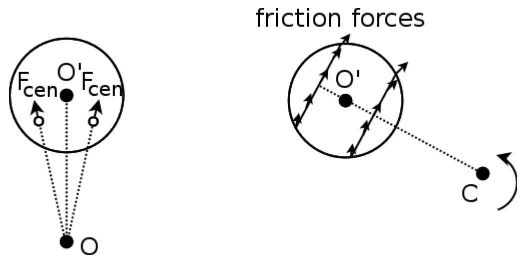


Figure 11.7: Centrifugal and friction forces

rotation (ICoR) which we call point C. Points on the rigid body that have equal speeds are equidistant from C. That is, they lie on an arc centered about C. Consider the set of points along one such arc — the Coriolis force on them is radially outwards (relative to the center C). Similar to the case of the centrifugal force which is also radial, the net torque about O' due to the Coriolis forces on pairs of points that are mirror images about the line CO' is zero — implying that the net torque about O' due to the Coriolis force on the entire disk is zero.

To analyze the torque due to friction, suppose that the ICoR is located at a finite distance from the center of the disk. Notice that every point on the disk experiences a friction force of the same magnitude that is perpendicular to the line joining C and it (right diagram of Fig. 11.7). Now, consider two chords that are perpendicular to line CO' and are the same perpendicular distance away from O' (they lie on different sides of O'). Since the chord closer to C subtends a larger angle with respect to C, the friction forces along this chord generally make larger angles with the chord — implying that the torque about O' due to friction along this chord is smaller in magnitude than the torque on its counterpart. In the case of the right figure where the angular velocity of the disk in the rotating frame is anti-clockwise relative to C, the friction forces on the disk are clockwise relative to C such that every pair of corresponding chords experience a net frictional torque clockwise about O' — violating the claim that the angular velocity of the disk is constant. Therefore, for the disk to attain a steady angular velocity (i.e. no torque about O'), its instantaneous center of rotation must be located at infinity, such that all points on the disk experience identical friction forces pointing towards the same direction (i.e. the disk solely translates and does not rotate about its center). This implies that the angular velocity of the disk is zero in the rotating frame and is thus ω anti-clockwise in the lab frame.

11. Straight Line on Carousel**

Consider a coordinate system fixed to the carousel and define the X and Y axes of this rotating frame to coincide with the x and y axes at time $t = 0$. By Eq. (11.17), the equation of motion of the particle in this frame is

$$m\begin{pmatrix}\ddot{X}\\ \ddot{Y}\\ \ddot{Z}\end{pmatrix} = m\omega^2\boldsymbol{r} + 2m\begin{pmatrix}\dot{X}\\ \dot{Y}\\ 0\end{pmatrix} \times \begin{pmatrix}0\\ 0\\ \omega\end{pmatrix},$$

$$m\begin{pmatrix}\ddot{X}\\ \ddot{Y}\\ \ddot{Z}\end{pmatrix} = m\omega^2\begin{pmatrix}X\\ Y\\ 0\end{pmatrix} + 2m\begin{pmatrix}\dot{Y}\omega\\ -\dot{X}\omega\\ 0\end{pmatrix},$$

$$\ddot{X} = \omega^2 X + 2\omega\dot{Y},$$
$$\ddot{Y} = \omega^2 Y - 2\omega\dot{X}.$$

Defining a new complex variable $\eta = X + iY$ and adding the product of the second equation with i to the first,

$$\ddot{\eta} + 2i\omega\dot{\eta} - \omega^2\eta = 0.$$

Guessing a solution of the form $\eta = \eta_0 e^{\alpha t}$,

$$\alpha^2 + 2i\omega\alpha - \omega^2 = 0$$
$$(\alpha + i\omega)^2 = 0,$$

and we only have one solution $\alpha = -i\omega$. However, there should be two independent solutions to accommodate two initial conditions (initial coordinates and velocities). Similar to the case of critical damping, the general solution for η is in fact

$$\eta = Ae^{-i\omega t} + Bte^{-i\omega t},$$

where A and B are complex constants. Substituting the initial condition $\eta = x_0 + iy_0$ at $t = 0$,

$$A = x_0 + iy_0.$$

Differentiating η with respect to time,

$$\dot{\eta} = -i\omega Ae^{-i\omega t} + Be^{-i\omega t} - i\omega Bte^{-i\omega t}.$$

Before enforcing the initial condition, we first determine $\dot{\eta}$ at $t = 0$ by determining the initial velocity of the particle in the rotating frame. Since

$$\boldsymbol{v} = \boldsymbol{v}_{rot} + \boldsymbol{\omega} \times \boldsymbol{r},$$

$$u_x = u_X - \omega y_0,$$

$$u_y = u_Y + \omega x_0,$$

where u_X and u_Y are the components of initial velocity in the rotating frame. Then, $\dot{\eta}$ at $t = 0$ is

$$u_X + iu_Y = u_x + \omega y_0 + i(u_y - \omega x_0).$$

Enforcing this initial condition in the solution we guessed for $\dot{\eta}$,

$$B = u_x + iu_y.$$

Therefore,

$$\eta = (x_0 + iy_0)e^{-i\omega t} + (u_x + iu_y)te^{-i\omega t}.$$

The physical X and Y can be determined by taking the real and imaginary components of the above respectively.

$$X = \text{Re}(\eta) = x_0 \cos \omega t + y_0 \sin \omega t + u_x t \cos \omega t + u_y t \sin \omega t,$$

$$Y = \text{Im}(\eta) = y_0 \cos \omega t - x_0 \sin \omega t + u_y t \cos \omega t - u_x t \sin \omega t.$$

Finally, when t is large, the oscillatory part of η with a constant amplitude pales in comparison with the other term whose amplitude scales with t. Thus, the former vanishes and η becomes

$$\eta \approx (u_x + iu_y)te^{-i\omega t} = \sqrt{u_x^2 + u_y^2}\, te^{-i(\omega t - \phi)},$$

where $\phi = \tan^{-1} \frac{u_y}{u_x}$. Then,

$$X = \sqrt{u_x^2 + u_y^2}\, t \cos(\omega t - \phi),$$

$$Y = -\sqrt{u_x^2 + u_y^2}\, t \sin(\omega t - \phi),$$

$$\implies X^2 + Y^2 = (u_x^2 + u_y^2)t^2.$$

Now, why is this shape of the form of a spiral? Notice that if the right-hand side is simply a constant, this equation describes a circle, with its radius being the square root of the right-hand side. However, in the current situation, the radius of the "circle" ($\sqrt{u_x^2 + u_y^2}\, t$) constantly increases while

X and Y oscillate between positive and negative values such that the particle rotates about the origin at an increasing radius — corresponding to a spiral.

12. Larmor's Theorem**

Define $f(\boldsymbol{r})$ to be the net force that a charge at position vector $\boldsymbol{r}$ experiences in the original experiment. Now, consider the set-up with the magnetic field and switch to a frame rotating with respect to the original frame at a certain angular velocity $\boldsymbol{\omega}$. The equation of motion of a charge in this frame is

$$\begin{aligned} m\boldsymbol{a}_{rot} &= f(\boldsymbol{r}) + q\boldsymbol{v}_{lab} \times \boldsymbol{B} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + 2m\boldsymbol{v}_{rot} \times \boldsymbol{\omega} \\ &= f(\boldsymbol{r}) + q(\boldsymbol{v}_{rot} + \boldsymbol{\omega} \times \boldsymbol{r}) \times \boldsymbol{B} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}) + 2m\boldsymbol{v}_{rot} \times \boldsymbol{\omega} \\ &= f(\boldsymbol{r}) + \boldsymbol{v}_{rot} \times (q\boldsymbol{B} + 2m\boldsymbol{\omega}) + q(\boldsymbol{\omega} \times \boldsymbol{r}) \times \boldsymbol{B} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}). \end{aligned}$$

Now, you might be confused about why the magnetic force is $q\boldsymbol{v}_{lab} \times \boldsymbol{B}$ and not $q\boldsymbol{v}_{rot} \times \boldsymbol{B}$. Well, this is because by definition, the magnetic force is a real force that a charge experiences and must thus be invariant across all frames in the context of classical mechanics. Therefore, we can compute it in the lab frame to obtain $q\boldsymbol{v}_{lab} \times \boldsymbol{B}$. Technically, we can also compute the Lorentz force on the charge in the rotating frame but it is no longer solely due to a magnetic force as a pure magnetic field in one frame generally transforms into an electric-cum-magnetic field (the magnetic field may not even be the same) in another frame. In fact, we usually make the above arguments in the reverse direction — by imposing the condition that the real Lorentz force on the charge must vary in a certain manner across different frames (we are no longer in the regime of classical mechanics where forces are necessarily invariant), we can compute the transformations of electromagnetic fields. Moving on, notice that if we choose

$$\boldsymbol{\omega} = -\frac{q\boldsymbol{B}}{2m},$$

the second term vanishes. Furthermore, the third and fourth terms will be second order in B^2 and will be negligible for small B when compared to $f(\boldsymbol{r})$. The equation of motion then reduces to

$$m\boldsymbol{a}_{rot} = f(\boldsymbol{r}),$$

which is simply the equation of motion in the original experiment! Remembering that the above quantities are observed in the rotating frame, the entire set-up must evolve in the same manner as the first experiment, while precessing at $\boldsymbol{\omega} = -\frac{q\boldsymbol{B}}{2m}$.

13. Sniping South**

Adopting the same axes definitions as in Fig. 11.5, the displacement of the bullet at time t is given by Eq. (11.22) as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} v_0 t \\ 0 \\ -\frac{gt^2}{2} \end{pmatrix} + t^2 \begin{pmatrix} v_0 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} \omega \sin\theta \\ 0 \\ \omega\cos\theta \end{pmatrix} + \frac{t^3}{3} \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \times \begin{pmatrix} \omega\sin\theta \\ 0 \\ \omega\cos\theta \end{pmatrix}$$
$$= \begin{pmatrix} v_0 t \\ -v_0 t^2 w \cos\theta - \frac{gt^3}{3}\omega\sin\theta \\ -\frac{gt^2}{2} \end{pmatrix}.$$

The only non-trivial solution to the y-component of the displacement being zero is

$$t = -\frac{3v_0}{g}\cot\theta.$$

Substituting this expression into the x-component,

$$x = -\frac{3v_0^2}{g}\cot\theta.$$

The southwards distance is hence $\frac{3v_0^2}{g}\cot\theta$.

This page intentionally left blank

Chapter 12

Lagrangian Mechanics

In the past few chapters, we have examined the quintessence of Newtonian mechanics. However, in this chapter, we shall adopt a completely different approach to dynamics via a general principle known as Hamilton's principle. In analytical mechanics, dynamics problems are reduced to writing down a function known as the Lagrangian, performing a few differentiations and then solving the ensuing differential equations. In fact, Lagrange prided himself on not including a single diagram in his treatise! We will delve directly into the heart of this topic, so please bear with the abrupt jump.

12.1 Action and Hamilton's Principle

The action S of a system along an evolutionary path $\boldsymbol{q}(t)$ between a starting state at time t_1 and an ending state at time t_2 is defined as

$$S = \int_{t_1}^{t_2} \mathcal{L}(q_1, q_2, \ldots, q_n, \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n, t)dt. \tag{12.1}$$

$\mathcal{L}$ is the Lagrangian and is a function given by

$$\mathcal{L} = T - V, \tag{12.2}$$

where T and V are the instantaneous kinetic and potential energies of the system. This seems like an odd combination but some analogies will be revealed soon. q_1, q_2,..., q_n are generalized coordinates which uniquely describe the state of a system. As their nomenclature implies, they are indeed quite all-encompassing as they can represent translational coordinates such as x, angular coordinates such as θ and even time-varying coordinates, such as those of a rotating coordinate system. For brevity, they are collectively represented as $\boldsymbol{q}$ which can be seen as an n-dimensional vector. Note that this vector has little physical meaning — evident from the fact that even the units of each q_i can be different. $\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n$ are the time derivatives of the generalized coordinates and are known as generalized velocities — collectively denoted by $\dot{\boldsymbol{q}}$.

Next, we move on to the crucial pillar of Lagrangian mechanics. **Hamilton's principle** states that the path $\boldsymbol{q}(t)$ taken by a system between times t_1 and t_2 is one that results in a stationary value of the action S. Notice that S is dependent on the entire function $\boldsymbol{q}(t)$ and is hence, known as a functional. To determine the condition that results in a stationary value of S, we turn to the calculus of variations.

12.2 Calculus of Variations

Firstly, we quantify what a stationary value of a functional actually means. Suppose that $\boldsymbol{q}_0(t)$ is the desired path that results in a stationary value of the functional

$$S = \int_{t_1}^{t_2} \mathcal{L}[\boldsymbol{q}(t)]dt.$$

Then, functions $\boldsymbol{q}(t)$'s in the immediate vicinity of $\boldsymbol{q}_0(t)$ result in values of S adopting deviations of second-order and above from the value of S_0 produced by $\boldsymbol{q}_0(t)$ for any possible deviation. Let us consider the case of a functional that involves only a single coordinate $x(t)$, its derivative $\dot{x}$ and t (which is just a variable that is not necessarily time).

$$S[x(t)] = \int_{t_1}^{t_2} \mathcal{L}(x, \dot{x}, t)dt.$$

Remember that the end points $x(t_1)$ and $x(t_2)$ are fixed at certain values, as we have predetermined initial and final states.[1] Usually if $f(x)$ were a function of x, we would determine points that yield $\frac{df}{dx} = 0$ to find stationary points. Since it is impossible to differentiate a variable with respect to an entire function, consider the substitution of

$$x(t) = x_0(t) + \alpha\eta(t)$$
$$\dot{x}(t) = \dot{x}_0(t) + \alpha\dot{\eta}(t),$$

[1]Note that the final state at time t_2 is not really well-specified. For example, if you are analyzing the motion of a simple projectile, there is no clear point in time where the motion ends. However, Hamilton's principle states that whatever time you choose as the final state, the actual path taken by the ball between time t_1 and this final time extremizes the action.

where α is a constant and $\eta(t)$ is any arbitrary function with $\eta(t_1) = \eta(t_2) = 0$ such that the endpoints of the path are kept fixed at $x(t_1) = x_0(t_1)$ and $x(t_2) = x_0(t_2)$. Together, $\alpha\eta$ is known as a variation of $x(t)$, denoted by δx. Then, S is now a function of α and we can consider its derivative with respect to α. In order for S to be stationary, we require $\frac{\partial S}{\partial \alpha} = 0$.

$$\frac{\partial S}{\partial \alpha} = \frac{\partial}{\partial \alpha}\int_{t_1}^{t_2} \mathcal{L} dt.$$

As $\mathcal{L}$ is not integrated with respect to α, we can bring the partial derivative into the integral such that it becomes a total derivative.

$$\begin{aligned}\frac{\partial S}{\partial \alpha} &= \int_{t_1}^{t_2} \frac{d\mathcal{L}}{d\alpha} dt \\ &= \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial x}\frac{\partial x}{\partial \alpha} dt + \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \dot{x}}\frac{\partial \dot{x}}{\partial \alpha} dt,\end{aligned}$$

where the term $\int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial t}\frac{\partial t}{\partial \alpha} dt = 0$ has been excluded as α is a time-independent constant. Now, note that

$$\begin{aligned}\frac{\partial x}{\partial \alpha} &= \eta(t) \\ \frac{\partial \dot{x}}{\partial \alpha} &= \dot{\eta}(t).\end{aligned}$$

Then,

$$\frac{\partial S}{\partial \alpha} = \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial x}\eta dt + \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial \dot{x}}\dot{\eta} dt.$$

Integrating the second term by parts,

$$\frac{\partial S}{\partial \alpha} = \int_{t_1}^{t_2} \frac{\partial \mathcal{L}}{\partial x}\eta dt + \frac{\partial \mathcal{L}}{\partial \dot{x}}\eta\Big|_{t_1}^{t_2} - \int_{t_1}^{t_2} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right)\eta dt.$$

Since $\eta(t_1) = \eta(t_2) = 0$,

$$\int_{t_1}^{t_2} \left(\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial x}\right)\eta dt = 0$$

in order for $\frac{\partial S}{\partial \alpha} = 0$. By the fundamental lemma of the calculus of variations, if

$$\int_{t_1}^{t_2} M(x,t)\eta(t) dt = 0$$

for all $\eta(t)$'s that are twice continuously differentiable, then

$$M(x,t) = 0.$$

Applied to the situation at hand where η is an arbitrary function, it implies that

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) = \frac{\partial \mathcal{L}}{\partial x}$$

for S to be extremized. This is known as the Euler-Lagrange equation (we will refer to this as the E-L equation for the sake of convenience). Another way to see why this must be true without the use of the fundamental lemma is to imagine choosing a $\eta(t)$ such that it has the same sign at all times with $(\frac{\partial \mathcal{L}}{\partial x} - \frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{x}})$ (it could be this expression scaled down by a certain factor). If the latter expression were not zero, the integrand would be positive at all times and, hence, cause the integral to be positive — leading to a contradiction. For functions that depend on n coordinates q_1, q_2,..., q_n, the above can be easily generalized by letting

$$q_i(t) = q_i^0(t) + \alpha_i \eta_i(t),$$

where $q_i^0(t)$ is the correct function for coordinate q_i and $\alpha_i \eta_i(t)$ is a variation along this coordinate. Then one can take

$$\frac{\partial S}{\partial \alpha_i} = 0$$

for all $1 \leq i \leq n$. Since adding a variation to the coordinates other than q_k does not result in a change in the partial derivative of the Lagrangian with respect to α_k, we can use the previous result to deduce that the overall condition for $S(\boldsymbol{q}(t))$ to undertake a stationary value is

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) = \frac{\partial \mathcal{L}}{\partial q_i}$$

for all $1 \leq i \leq n$. We can express this in a rather succinct form

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}},$$

where a derivative of a function y with respect to $\boldsymbol{q}$ means

$$\frac{\partial y}{\partial \boldsymbol{q}} = \begin{pmatrix} \frac{\partial y}{\partial q_1} \\ \frac{\partial y}{\partial q_2} \\ \vdots \\ \frac{\partial y}{\partial q_n} \end{pmatrix}.$$

12.3 Equations of Motion

Now that we have understood the requirements for a function to extremize a functional, let us return to our original question regarding the action that depends on the path taken by a system

$$S = \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) dt.$$

To attain a stationary value for S, the E-L equations require that

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}}. \tag{12.3}$$

The desired $\boldsymbol{q}(t)$ that extremizes the action is known as an **extremal**. Let us see how the Newtonian equations of motion in Cartesian coordinates can be "recovered" in the case of a single particle. In Cartesian coordinates,

$$\mathcal{L} = T - V = \frac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - V.$$

Then,

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) &= \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}} \\ \implies m\ddot{x} &= -\frac{\partial V}{\partial x}, \\ m\ddot{y} &= -\frac{\partial V}{\partial y}, \\ m\ddot{z} &= -\frac{\partial V}{\partial z}, \\ \implies m\ddot{\boldsymbol{r}} &= -\frac{\partial V}{\partial \boldsymbol{r}}. \end{aligned}$$

Notice that $-\frac{\partial V}{\partial \boldsymbol{r}}$ is simply the net conservative external force on the particle. Therefore, we retrieve the equation

$$\boldsymbol{F} = m\boldsymbol{a}.$$

For a system of particles, our generalized coordinates could be the union of all (x_i, y_i, z_i) for each particle i and the $\boldsymbol{F} = m\boldsymbol{a}$ equation would hold for each particle. Newton's second law is therefore equivalent to Hamilton's principle applied in Cartesian coordinates!

From the above example, it should be clear that there are analogies between the Newtonian momentum and forces with the terms in the E-L equations. $p_i \equiv \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$ is known as the **canonical momentum conjugate to coordinate** q_i, while $Q_i \equiv \frac{\partial \mathcal{L}}{\partial q_i}$ is known as the **generalized force associated with coordinate** q_i.

What about non-conservative forces? The fact is that from experience, all of the fundamental forces that we know today can be formulated in terms of potentials. Non-conservative forces, such as the viscous drag force which originates from the bombardment of fluid particles on an object, are not fundamental forces. Therefore, the Lagrangian formulation is still general at the most fundamental level. However, if we utterly insist on including non-conservative forces, we can "cheat" a little and add an appropriate expression on the right-hand side of Eq. (12.3) to represent a non-conservative force.

12.3.1 *Different Coordinate Systems*

The Lagrangian may not seem particularly enlightening right now but its utility really shines when it comes to different coordinate systems. If we were to use Newton's laws, which are only valid in an inertial frame, in a rotating frame for example, we would have to modify the laws of motion to include "fictitious forces." However, by the Lagrangian formulation, the fundamental Hamilton's principle holds in all frames. Concomitantly, the E-L equations hold for various coordinate systems — we simply have to express T and V in terms of the different coordinates. This can be seen in two ways: physically and mathematically. Physically, the path of an extremal should not depend on the frame of reference. For example, the shortest path between two points is a straight line, regardless of the frame it is viewed from. Mathematically, we can show that if the E-L equations hold for a set of n coordinates $\boldsymbol{x}(t)$ (and we already know that it holds for the Cartesian coordinate system), it must hold for the same Lagrangian in another set of N coordinates[2] $\boldsymbol{q}(t)$ that is given by

$$\begin{aligned} q_i(t) &= f_i(\boldsymbol{x}, t) \\ \implies \dot{q}_i(t) &= \dot{f}_i(\boldsymbol{x}, \dot{\boldsymbol{x}}, t) \end{aligned} \tag{12.4}$$

for all $1 \leq i \leq N$. That is, each transformed coordinate is a function of the previous coordinates and time t. The following proof is just a formality and can be skipped by readers who simply want to learn how the Lagrangian can be applied.

[2] N is not necessarily equal to n.

Returning to the problem, we wish to prove that

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}}$$

given that

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{x}}}\right) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{x}}.$$

Consider

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial q_k} &= \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial x_i} \cdot \frac{\partial x_i}{\partial q_k} + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \cdot \frac{\partial \dot{x}_i}{\partial q_k} \\ &= \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right) \cdot \frac{\partial x_i}{\partial q_k} + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \cdot \frac{\partial \dot{x}_i}{\partial q_k}.\end{aligned}$$

For the final term on the right, observe that

$$\begin{aligned}\frac{\partial \dot{x}_i}{\partial q_k} &= \frac{\partial}{\partial q_k}\left(\sum_{j=1}^{N} \frac{\partial x_i}{\partial q_j} \cdot \dot{q}_j + \frac{\partial x_i}{\partial t}\right) \\ &= \sum_{j=1}^{N} \frac{\partial}{\partial q_j}\left(\frac{\partial x_i}{\partial q_k}\right) \cdot \dot{q}_j + \frac{\partial}{\partial t}\left(\frac{\partial x_i}{\partial q_k}\right) \\ &= \frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right).\end{aligned}$$

Note that we have used the fact that partial derivatives are interchangeable in writing the second inequality. Therefore,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial q_k} &= \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i}\right) \cdot \frac{\partial x_i}{\partial q_k} + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \cdot \frac{d}{dt}\left(\frac{\partial x_i}{\partial q_k}\right) \\ &= \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \cdot \frac{\partial x_i}{\partial q_k}\right).\end{aligned}$$

As our objective is to show that this is equal to $\frac{d}{dt}(\frac{\partial \mathcal{L}}{\partial \dot{q}_k})$, we simply have to prove that

$$\frac{\partial x_i}{\partial q_k} = \frac{\partial \dot{x}_i}{\partial \dot{q}_k}.$$

Observe that

$$\dot{x}_i = \sum_{j=1}^{N} \frac{\partial x_i}{\partial q_j} \cdot \dot{q}_j + \frac{\partial x_i}{\partial t}$$

$$\implies \frac{\partial \dot{x}_i}{\partial \dot{q}_k} = \frac{\partial x_i}{\partial q_k}.$$

Thus,

$$\frac{\partial \mathcal{L}}{\partial q_k} = \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} \cdot \frac{\partial \dot{x}_i}{\partial \dot{q}_k}\right) = \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial q_k}\right)$$

for all $1 \leq k \leq N$. This implies that

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\boldsymbol{q}}}\right) = \frac{\partial \mathcal{L}}{\partial \boldsymbol{q}}.$$

If the E-L equations are valid for one set of coordinates (for example, Cartesian coordinates), they are also valid for all other coordinates of the form given by Eq. (12.4). A direct corollary of this is that frames related by Galilean transformations are equivalent as the coordinate transformations are of the form $q_i' = q_i + v_i t$ for some constant v_i.

Polar Coordinates

The Lagrangian really distinguishes itself in finding the $\boldsymbol{F} = m\boldsymbol{a}$ equations of a single particle in polar coordinates — whose transformations from Cartesian coordinates evidently obey Eq. (12.4). Our objective is to express T in terms of polar coordinates. To do so, we have to determine $|\frac{d\boldsymbol{r}}{dt}|^2$ where $\boldsymbol{r}$ is the position vector of the particle. This can be written as

$$\left|\frac{d\boldsymbol{r}}{dt}\right|^2 = \frac{d\boldsymbol{r} \cdot d\boldsymbol{r}}{(dt)^2} = \frac{(dr_1)^2}{(dt)^2} + \frac{(dr_2)^2}{(dt)^2} + \frac{(dr_3)^2}{(dt)^2},$$

where dr_1, dr_2 and dr_3 are three perpendicular infinitesimal length segments (the usual infinitesimal quantities we integrate over). Let us apply this to an example.

Spherical Coordinates

The infinitesimal length segments in spherical coordinates are $dr_1 = dr$, $dr_2 = r\sin\theta d\phi$ and $dr_3 = rd\theta$. Hence, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}m\left(\dot{r}^2 + r^2\sin^2\theta\dot{\phi}^2 + r^2\dot{\theta}^2\right) - V.$$

We wish to evaluate how the change of these coordinates relates to the conservative forces along the unit vectors in spherical coordinates[3] which are

$$\begin{pmatrix} F_r \\ F_\phi \\ F_\theta \end{pmatrix} = \begin{pmatrix} -\frac{\partial V}{\partial r_1} \\ -\frac{\partial V}{\partial r_2} \\ -\frac{\partial V}{\partial r_3} \end{pmatrix} = \begin{pmatrix} -\frac{\partial V}{\partial r} \\ -\frac{\partial V}{r\sin\theta\partial\phi} \\ -\frac{\partial V}{r\partial\theta} \end{pmatrix}. \tag{12.5}$$

Applying the E-L equations ($\frac{\partial\mathcal{L}}{\partial\boldsymbol{q}} = \frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{q}}})$) with respect to generalized coordinates r, ϕ and θ respectively,

$$-\frac{\partial V}{\partial r} + mr\sin^2\theta\dot{\phi}^2 + mr\dot{\theta}^2 = \frac{d}{dt}(m\dot{r}) = m\ddot{r},$$

$$-\frac{\partial V}{\partial\phi} = \frac{d}{dt}\left(mr^2\sin^2\theta\dot{\phi}\right) = 2mr\sin^2\theta\dot{r}\dot{\phi} + 2mr^2\sin\theta\cos\theta\dot{\phi}\dot{\theta} + mr^2\sin^2\theta\ddot{\phi},$$

$$-\frac{\partial V}{\partial\theta} + mr^2\sin\theta\cos\theta\dot{\phi}^2 = \frac{d}{dt}\left(mr^2\dot{\theta}\right) = 2mr\dot{r}\dot{\theta} + mr^2\ddot{\theta}.$$

Dividing the second and third equations by $r\sin\theta$ and r respectively and applying Eq. (12.5),

$$\begin{aligned} F_r &= m(\ddot{r} - r\sin^2\theta\dot{\phi}^2 - r\dot{\theta}^2), \\ F_\phi &= m(2\sin\theta\dot{r}\dot{\phi} + 2r\cos\theta\dot{\phi}\dot{\theta} + r\sin\theta\ddot{\phi}), \\ F_\theta &= m(2\dot{r}\dot{\theta} - r\sin\theta\cos\theta\dot{\phi}^2 + r\ddot{\theta}). \end{aligned}$$

We have derived such a complicated version of $\boldsymbol{F} = m\boldsymbol{a}$ without the use of any vectors! This is the slickness of the Lagrangian.

Rotating Coordinate Systems

Consider a rotating coordinate system X, Y, Z that is rotating at a constant anti-clockwise angular velocity ω around the z-axis, of a fixed coordinate

[3]This is just the negative gradient of V in spherical coordinates (the definition of a conservative force).

frame x, y, z which coincides with the rotating frame at $t = 0$. Then, the x, y, z coordinates can be expressed in terms of X, Y and Z as

$$x = X\cos\omega t - Y\sin\omega t,$$
$$y = X\sin\omega t + Y\cos\omega t,$$
$$z = Z.$$

The first two equations can be swiftly obtained from applying the rotation matrix $\begin{pmatrix}\cos\omega t & -\sin\omega t\\ \sin\omega t & \cos\omega t\end{pmatrix}$ to $\begin{pmatrix}X\\ Y\end{pmatrix}$. In doing so, be cautious that the x and y-axes rotate at ω **clockwise** relative to the X and Y-axes.

$$\dot{x} = \dot{X}\cos\omega t - \omega X\sin\omega t - \dot{Y}\sin\omega t - \omega Y\cos\omega t,$$
$$\dot{y} = \dot{X}\sin\omega t + \omega X\cos\omega t + \dot{Y}\cos\omega t - \omega Y\sin\omega t,$$
$$\dot{z} = \dot{Z}.$$

The Lagrangian in terms of the fixed coordinate frame is

$$\mathcal{L} = \frac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2\right) - V.$$

This can be expressed in terms of the rotating coordinates as

$$\mathcal{L} = \frac{1}{2}m\left[\dot{X}^2 + \dot{Y}^2 + \dot{Z}^2 + 2\omega X\dot{Y} - 2\omega\dot{X}Y + \omega^2(X^2 + Y^2)\right] - V.$$

Applying the E-L equations ($\frac{\partial\mathcal{L}}{\partial\boldsymbol{q}} = \frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{q}}})$) with respect to X, Y and Z yields

$$m\omega\dot{Y} - \frac{\partial V}{\partial X} + m\omega^2 X = m\ddot{X} - m\omega\dot{Y},$$
$$-m\omega\dot{X} - \frac{\partial V}{\partial Y} + m\omega^2 Y = m\ddot{Y} + m\omega\dot{X},$$
$$-\frac{\partial V}{\partial Z} = m\ddot{Z}.$$

Note that $(-\frac{\partial V}{\partial X}, -\frac{\partial V}{\partial Y}, -\frac{\partial V}{\partial Z})$, where the unit vectors are along the axes of the rotating frame, represents the real conservative force $\boldsymbol{F}$ which is invariant across all frames (as it is a physical vector). Let $\boldsymbol{r}$ be the position vector of the particle which is independent of the frame of reference. Furthermore, denote $\boldsymbol{v}_{rot} = (\dot{X}, \dot{Y}, \dot{Z})$ and $\boldsymbol{a}_{rot} = (\ddot{X}, \ddot{Y}, \ddot{Z})$ as the particle's velocity and acceleration, as observed in the rotating frame, respectively. Summarizing the above equations in terms of vectors would yield

$$m\boldsymbol{a}_{rot} = \boldsymbol{F} - 2m\boldsymbol{\omega} \times \boldsymbol{v}_{rot} - m\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \boldsymbol{r}).$$

Keep in mind that $\boldsymbol{\omega}$ is the angular velocity of the rotating frame with respect to the lab frame and points in the positive z-direction.

12.3.2 *Additional Time Derivatives*

The Lagrangian describing the evolution of a system is in fact not unique. Try adding a total time derivative of a function of generalized coordinates and time $\frac{df(\boldsymbol{q},t)}{dt}$ to the original Lagrangian of a system. Note that $\frac{df}{dt}$ could simply be a constant too. The action between times t_1 and t_2 becomes

$$S' = \int_{t_1}^{t_2} \left(\mathcal{L} + \frac{df(\boldsymbol{q},t)}{dt} \right) dt = \int_{t_1}^{t_2} \mathcal{L} dt + f(\boldsymbol{q}_2, t_2) - f(\boldsymbol{q}_1, t_1),$$

where $\boldsymbol{q}_2$ and $\boldsymbol{q}_1$ are the final and initial generalized coordinates respectively. Those additional terms on the right are fixed and do not vary — the equations of motion of this new system thus do not differ from the original one. This implies that total time derivatives can simply be discarded from the Lagrangian of a system as they are inconsequential — a neat trick in tidying up the Lagrangian.

Problem: A particle of mass m is attached to a wall via a massless spring with spring constant k. In this one-dimensional problem, the x-coordinate of the wall is constrained to obey $X(t) = A\cos\omega t$ where $A > 0$ is a constant while the particle and the origin lie on opposite sides of the wall. Let x denote the additional displacement of the particle from the wall, beyond the rest length of the spring, in the positive x-direction. Show that

$$\ddot{x} + \omega_0^2 x = B\cos\omega t,$$

where $\omega_0 = \sqrt{\frac{k}{m}}$ is the natural frequency of this system and B is a constant.

The velocity of the particle is $\dot{X} + \dot{x} = \dot{x} - A\omega\sin\omega t$. The Lagrangian of the particle is thus

$$\mathcal{L} = \frac{1}{2}m(\dot{x} - A\omega\sin\omega t)^2 - \frac{1}{2}kx^2.$$

We can discard the $\frac{1}{2}mA^2\omega^2\sin^2\omega t$ term since it is a total time derivative, to obtain

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 - mA\omega\sin\omega t\dot{x} - \frac{1}{2}kx^2.$$

Applying the E-L equation $\left(\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) = \frac{\partial \mathcal{L}}{\partial x}\right)$,

$$m\ddot{x} - mA\omega^2 \cos \omega t = -kx$$

$$\ddot{x} + \omega_0^2 x = A\omega^2 \cos \omega t,$$

as required by the problem. Observe that this equation of motion describes a forced oscillation whose general solution can be determined by methods introduced in the chapter on oscillations.

12.4 Systems with Constraints

In many mechanics problems, there are constraints imposed on a system. For example, a falling object cannot penetrate the ground and the length of a rigid rod remains constant. Constraints can be holonomic or non-holonomic. A holonomic constraint can be expressed algebraically in the following form:

$$f(\boldsymbol{q}, t) = 0.$$

An example of a holonomic constraint would be that of rigid bodies — the preservation of relative distances can be expressed in terms of the generalized coordinates. Constraints that cannot fulfil the above criterion are termed non-holonomic — they usually involve time derivatives of generalized coordinates, and inequalities instead of equalities. Note that the constraint on a rigid body to roll without slipping along a single direction, despite its appearance, is holonomic as the time derivatives can be easily removed via integration.

The Lagrangian formulation can only be conveniently applied to holonomic systems. The two general methods in doing so will be elaborated in the following section.

But first, there is an often simpler approach to determine whether a system is holonomic or non-holonomic, instead of expressing the constraints algebraically. Define the degrees of freedom (DOFs) of a system as the number of coordinates of a system that can be varied independently of all other coordinates (i.e. keep the other coordinates fixed). Let there be N particles in a system. Then, an unconstrained system has $3N$ DOFs and requires $3N$ coordinates to specify a state. In holonomic systems, the DOFs of a system are equal to the number of independent coordinates required to uniquely specify a state of a system. This is due to m holonomic constraints reducing the DOFs by m. From a mathematical perspective, holonomic constraints furnish m additional equations in terms of the generalized coordinates and time such that the $3N$ coordinates can be expressed in terms of $(3N - m)$

independent ones. A non-holonomic system, naturally, does not fulfil the criterion that the DOFs are equal to the number of independent coordinates needed to uniquely define a state.

Example of a Non-holonomic System

One might wonder how such an intuitive requirement can be violated. Consider the classic example of a sphere on a two-dimensional table. It is not allowed to translate or rotate about an axis in the vertical direction. Furthermore, it is constrained such that it cannot slip. Then, the sphere only has 2 DOFs — rolling in the x and y directions along the table. Now you might think that one simply needs two coordinates such as the x and y-coordinates of the center of the sphere to specify its state. Let us perform an experiment to convince you otherwise. Place the sphere at the origin initially and paste a sticker on top of it. Then, consider the following series of movements.

(1) Let the circumference of the sphere be C. Then, roll the sphere to coordinate $(C, 0)$. The sticker is on top of the sphere at this juncture. Then, roll it to coordinate (C, C). The sticker is still on top.
(2) Roll the sphere directly along the diagonal from the origin to (C, C). The sticker is no longer on top of the sphere!

In this system, 4 coordinates are in fact necessary to describe a configuration of the sphere! The non-holonomic constraint originates from the requirement of rolling without slipping in **two dimensions** such that the velocities in two directions are inextricably coupled in an equation which cannot be trivially integrated to remove the time derivatives. Anyway, non-holonomic systems will not be considered in this chapter and our analysis of such systems ends here.

Hamilton's Principle in Holonomic Systems

We cannot directly apply the previous results — namely the E-L equations to holonomic systems as the variations may not be consistent with the constraints. However, Hamilton's principle still holds and our modified objective is to determine a path that extremizes the action, while obeying the constraints. There are then two approaches that we can take.

Firstly, we can use m constraint equations to solve for $(3N - m)$ independent coordinates which can be used to define the state of a system. Often, we can even directly define coordinates that satisfy the constraints. The variations then naturally obey the constraints — the E-L equations can consequently be directly applied to this judiciously chosen set of independent coordinates. The greatest benefit of this approach is that the forces of constraint, such as normal forces, are completely ignored. As this method is rather straightforward, let us consider two examples to summarize the concepts so far.

Problem: A block of mass m lies on a frictionless inclined plane of mass M and angle of inclination θ. If there is no friction between the plane and the ground, determine the acceleration of the plane.

Define x to be the horizontal coordinate of the highest tip of the plane. Take the vertical coordinate to be zero at this tip. Then, define s to be the distance between the mass m and this tip (we assume that the gradient of the slope is negative in the positive x-direction). This definition ensures that m satisfies the constraint of remaining on the plane. The coordinates of m are thus $(x + s\cos\theta, -s\sin\theta)$. The velocity of m is $(\dot{x} + \dot{s}\cos\theta, -\dot{s}\sin\theta)$. The Lagrangian of the combined system comprising m and M is

$$\begin{aligned}\mathcal{L} &= T - V \\ &= \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m\left[(\dot{x} + \dot{s}\cos\theta)^2 + \dot{s}^2\sin^2\theta\right] + mgs\sin\theta \\ &= \frac{1}{2}(M+m)\dot{x}^2 + \frac{1}{2}m\dot{s}^2 + m\dot{x}\dot{s}\cos\theta + mgs\sin\theta.\end{aligned}$$

Applying the E-L equation ($\frac{d}{dt}(\frac{\partial\mathcal{L}}{\partial\dot{\boldsymbol{q}}}) = \frac{\partial\mathcal{L}}{\partial\boldsymbol{q}}$) with respect to x and s,

$$\begin{aligned}(M+m)\ddot{x} + m\ddot{s}\cos\theta &= 0, \\ m\ddot{s} + m\ddot{x}\cos\theta &= mg\sin\theta.\end{aligned}$$

Solving for $\ddot{x}$,

$$\ddot{x} = -\frac{mg\sin\theta\cos\theta}{M + m\sin^2\theta}.$$

Problem: Consider a mass m undergoing uniform circular motion at angular velocity ω in a fixed, hollow inverted cone of half angle θ. Determine the equations of motion of the particle and hence find l_0, the distance of the particle from the apex such that this uniform circular motion can occur. Then, show that this equilibrium is stable and determine the frequency of small oscillations Ω about l_0 when the particle is slightly perturbed in the direction joining the apex and the particle.

Let the distance of the particle to the apex be l and the azimuthal angular coordinate of the particle be ϕ. Then, $l\sin\theta\dot{\phi}$ is the azimuthal velocity while

$\dot{l}$ is the other component of velocity along the surface of the cone. Hence, the Lagrangian of this particle is

$$\mathcal{L} = \frac{1}{2}m(\dot{l}^2 + l^2 \sin^2\theta\dot{\phi}^2) - mgl\cos\theta.$$

Applying the E-L equation with respect to l and ϕ,

$$m\ddot{l} = ml\sin^2\theta\dot{\phi}^2 - mg\cos\theta,$$

$$\frac{d}{dt}(ml^2\sin^2\theta\dot{\phi}) = 0 \implies ml^2\sin^2\theta\dot{\phi} = L,$$

for some constant L. The Newtonian interpretation of this is simply the component of the angular momentum of the particle along the symmetrical axis of the cone. The above are the equations of motion for the particle. To determine l_0, we set $\ddot{l} = 0$ and $\dot{\phi} = \omega$. Then,

$$l_0 = \frac{g\cos\theta}{\omega^2\sin^2\theta}.$$

Now, consider a perturbation from l_0 such that $l = l_0 + \varepsilon$ for some infinitesimal deviation ε. We first rewrite the first equation of motion strictly in terms of a single variable l with the help of L:

$$m\ddot{l} = \frac{L^2}{ml^3\sin^2\theta} - mg\cos\theta.$$

Keep in mind that

$$\frac{L^2}{ml_0^3\sin^2\theta} - mg\cos\theta = 0, \tag{12.6}$$

as this will be used later. We then substitute $l = l_0 + \varepsilon$ into the previous equation to get

$$m\ddot{\varepsilon} = \frac{L^2}{ml_0^3\sin^2\theta\left(1 + \frac{\varepsilon}{l_0}\right)^3} - mg\cos\theta.$$

Using the first-order binomial expansion $(1+x)^n \approx 1 + nx$,

$$m\ddot{\varepsilon} = \frac{L^2}{ml_0^3\sin^2\theta}\left(1 - 3\frac{\varepsilon}{l_0}\right) - mg\cos\theta.$$

Applying Eq. (12.6) yields

$$\ddot{\varepsilon} = -\frac{3L^2}{m^2l_0^4\sin^2\theta}\varepsilon.$$

Note that such cancellations always occur in perturbation problems. The equilibrium is stable as the particle will undergo simple harmonic motion about l_0, after a slight deviation, at a frequency

$$\Omega = \frac{\sqrt{3}\,|L|}{ml_0^2 \sin\theta}.$$

Since $L = ml_0^2 \sin^2\theta\omega$,

$$\Omega = \sqrt{3}\sin\theta|\omega|.$$

Lagrange Multipliers

The second method entails directly finding variations that adhere to the constraints without changing the generalized coordinates. Mathematically, if there are m holonomic constraints and n generalized coordinates, we wish to extremize the functional

$$S = \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)dt$$

under m constraints, with the ith constraint being

$$f_i(\boldsymbol{q}, t) = 0.$$

Taking the total derivative of the above equation, we obtain the relationship between the variations of generalized coordinates that is imposed by the ith constraint.

$$\frac{\partial f_i}{\partial q_1}\delta q_1 + \frac{\partial f_i}{\partial q_2}\delta q_2 + \cdots + \frac{\partial f_i}{\partial q_n}\delta q_n = 0. \tag{12.7}$$

Now, consider the most general variation of the action which depends on $\boldsymbol{q}$ and $\dot{\boldsymbol{q}}$ (t cannot be varied):

$$\delta S = \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial q_j}\delta q_j + \frac{\partial \mathcal{L}}{\partial \dot{q}_j}\delta \dot{q}_j \right) dt.$$

The $\delta\dot{q}_j$ term for all j can be integrated by parts, while keeping in mind that $\delta q_j(t_1) = \delta q_j(t_2) = 0$ (as the end points must be fixed), to obtain

$$\delta S = \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j} \right) \right) \delta q_j dt.$$

Since the right-hand side of Eq. (12.7) is zero, we can add to the above equation Eq. (12.7), for each $1 \leq i \leq m$, multiplied by an arbitrary function of time $\lambda_i(t)$ without affecting the value of δS. These variables $\{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, collectively referred to as $\boldsymbol{\lambda}$, are known as **Lagrange Multipliers**.

$$\delta S = \int_{t_1}^{t_2} \sum_{j=1}^{n} \left(\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) + \sum_{i=1}^{m} \lambda_i(t) \frac{\partial f_i}{\partial q_j} \right) \delta q_j dt.$$

We can always find a $\boldsymbol{\lambda}$ such that the terms in the brackets equate to zero for $j = \{1, 2, \ldots, m\}$. Effectively, the first m variations, $\{\delta q_1, \delta q_2, \ldots, \delta q_m\}$, are expressed in terms of the other $(n - m)$ variations. Then,

$$\frac{\partial \mathcal{L}}{\partial q_j} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) + \sum_{i=1}^{m} \lambda_i(t) \frac{\partial f_i}{\partial q_j} = 0$$

for $1 \leq j \leq m$. Furthermore, the variations $\{\delta q_{m+1}, \delta q_{m+2}, \ldots, \delta q_n\}$ have been decoupled and thus can be varied independently. This is because the first m variations have been expressed in terms of them in a manner such that given $\{\delta q_{m+1}, \delta q_{m+2}, \ldots, \delta q_n\}$, one can always tweak $\{\delta q_1, \delta q_2, \ldots, \delta q_m\}$ such that they collectively satisfy the constraints. Since the last $(n - m)$ variations are free to take on any functions, the term in the brackets for each $m + 1 \leq j \leq n$ must also be zero for S to be stationary (by the fundamental lemma). All-in-all,

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = \sum_{i=1}^{m} \lambda_i(t) \frac{\partial f_i}{\partial q_j} \tag{12.8}$$

for all $1 \leq j \leq n$! These are the modified E-L equations for systems with holonomic constraints. There is a neat way to recapitulate the results derived.

In order to solve an extremization problem with holomonic constraints, we introduce m additional coordinates, $\lambda_i(t)$. Then, the above problem can be solved by finding the stationary values of the new action

$$S' = \int_{t_1}^{t_2} \left[\mathcal{L}(\boldsymbol{q}, \dot{\boldsymbol{q}}, t) + \sum_{i=1}^{m} \lambda_i(t) f_i(\boldsymbol{q}, t) \right] dt,$$

with $(n + m)$ generalized coordinates $\{\boldsymbol{q}, \boldsymbol{\lambda}\}$. The modified Lagrangian is

$$\mathcal{L}' = \mathcal{L} + \sum_{i=1}^{m} \lambda_i f_i.$$

To see why this is coherent with the previous results, apply the E-L equations with respect to a coordinate λ_i. Since $\mathcal{L}'$ is independent of $\dot{\lambda}_i$,

$$f_i(\boldsymbol{q}, t) = 0.$$

It can be seen that the constraints are enforced by introducing these additional terms. Therefore, we can vary $\boldsymbol{q}$ freely in a certain sense now as the extremals of S' will definitely obey the constraints. Furthermore, it is obvious that if S' is extremized, then S is also extremized (in a legal manner consistent with the constraints) as $\sum_{i=1}^{m} \lambda_i(t) f_i(\boldsymbol{q}, t) = 0$ when S' is extremized. Applying the E-L equations to $\mathcal{L}'$, with respect to q_j, yields

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right) - \frac{\partial \mathcal{L}}{\partial q_j} = \sum_{i=1}^{m} \lambda_i \frac{\partial f_i}{\partial q_j}.$$

Therefore, this method of modifying the Lagrangian is equivalent to the previous discussion and is valid. There are now $(n+m)$ E-L equations to solve for $(n+m)$ variables $\{\boldsymbol{q}, \boldsymbol{\lambda}\}$.

Finally, the term on the right-hand side of Eq. (12.8) has a physical meaning. If there were no constraints, the right-hand side should be zero. Then, this additional term must be due to the generalized force exerted by the constraints in this case!

$$Q_{j,\,constraint} = \sum_{i=1}^{m} \lambda_i \frac{\partial f_i}{\partial q_j}. \tag{12.9}$$

If q_j is a particular translational coordinate, $Q_{j,\,constraint}$ corresponds to the forces of constraint, such as static friction and normal force, along that coordinate! If q_j is an angular coordinate, one would have to modify Eq. (12.8) such that the denominator of the partial derivative $\frac{\partial \mathcal{L}}{\partial q_j}$ becomes an infinitesimal length segment (similar to our derivation of $\boldsymbol{F} = m\boldsymbol{a}$ in spherical coordinates) to obtain that particular component of the forces of constraint.

Problem: Let us consider the trivial example of calculating the tension in a simple pendulum. Let the fixed end of the string be at the origin and let the coordinates of the bob be (r, θ) where θ is the anti-clockwise angle between the string and the negative y-axis, that is pointing vertically downwards. There is a holonomic constraint $r = l$ where l is the length of the inextensible string. Then, the modified Lagrangian (we will still use $\mathcal{L}$ to represent it) is

$$\mathcal{L} = \frac{1}{2}m\left(\dot{r}^2 + r^2\dot{\theta}^2\right) + mgr\cos\theta + \lambda(r - l).$$

The E-L equation with respect to r is

$$m\ddot{r} = mr\dot{\theta}^2 + mg\cos\theta + \lambda.$$

Now applying the constraint $r = l$, $\ddot{r} = 0$,

$$\lambda = -mr\dot{\theta}^2 - mg\cos\theta.$$

The force of constraint along the r direction, which is simply the tension force on the bob due to the string in this case, is

$$T = Q_{r,\,constraint} = \lambda\frac{\partial(r-l)}{\partial r} = \lambda = -mr\dot{\theta}^2 - mg\cos\theta,$$

where the negative sign indicates that the direction of the force on the bob is radially inwards.

Problem: A point mass m initially rests on top of an immobile, frictionless circle of radius R. If it is given a slight displacement, determine the angle from the vertical at which it loses contact with the circle.

Although the constraint in this problem is technically that the mass cannot penetrate the circle (i.e. the radial coordinate of m with respect to the center satisfies $r \geq R$), the only relevant regime is where $r = R$. Therefore, we shall take $r = R$ as our holonomic constraint as we can identify the exact moment where it fails (i.e. when the normal force becomes zero). Define θ to be the angle that the position vector of m makes with the vertical axis, which is positive upwards. The modified Lagrangian is

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) - mgr\cos\theta + \lambda(r - R).$$

This Lagrangian is similar to that in the previous problem, with an additional negative sign in front of $mgr\cos\theta$ due to the differing definitions of θ. Then, the normal force N in this case can be obtained by substituting $(\pi - \theta)$ into the previous expression for tension, i.e.

$$N = \lambda\frac{\partial(r-R)}{\partial r} = -mr\dot{\theta}^2 + mg\cos\theta.$$

Next, we apply the E-L equation with respect to θ. In doing so, we can treat r as a constant ($r = R$) as the Lagrange multiplier is independent of θ such that its presence does not affect the equation of motion with respect to this coordinate.

$$g\sin\theta = R\ddot{\theta}.$$

Adopting the substitution $\ddot{\theta} = \frac{d(\dot{\theta}^2)}{2d\theta}$ and separating variables,

$$\int_0^{\dot{\theta}^2} d(\dot{\theta}^2) = \int_0^{\theta} \frac{2g}{R}\sin\theta d\theta$$

$$\dot{\theta}^2 = \frac{2g}{R}(1-\cos\theta).$$

Armed with this expression, we can determine the angle θ at which $N = 0$. When $N = -mr\dot{\theta}^2 + mg\cos\theta = 0$,

$$\dot{\theta}^2 = \frac{g}{R}\cos\theta$$

$$\frac{g}{R}\cos\theta = \frac{2g}{R}(1-\cos\theta)$$

$$\theta = \cos^{-1}\frac{2}{3}.$$

At this point in time, we are pining for the conservation of energy — a crucial component of Newtonian mechanics — as it would have drastically simplified the process in this problem.

12.5 Conserved Quantities and Symmetry*

The conservations of energy, momentum and angular momentum and more general forms of conserved quantities are actually direct consequences of Hamilton's principle. In fact, they are closely related to the symmetries of a system. This section is a formality in the sense that it simply proves that the Lagrangian formulation is coherent with several defining aspects of the Newtonian one. It is often much simpler to directly use the Newtonian formulation to obtain the conserved quantities.

12.5.1 *The Conservation of Energy*

If the Lagrangian is not an explicit function of t, we claim that the quantity

$$H \equiv \left(\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{q}_i\right) - \mathcal{L}$$

is conserved (time-independent). In a closed system, the Lagrangian is indeed time-independent as we should in principle be able to describe all states — past, present and future — in terms of the generalized coordinates and their first-order time derivatives.

H is known as the Hamiltonian of the system but is, most of the time, equivalent to the mechanical energy of the system. Since we will not be analyzing the Hamiltonian formulation here, we shall just refer to H as the mechanical energy of the system. To prove the claim above, consider the total time derivative of the Lagrangian.

$$\begin{aligned}\frac{d\mathcal{L}}{dt} &= \frac{\partial \mathcal{L}}{\partial t} + \sum_{i=1}^{n} \frac{\partial L}{\partial q_i} \cdot \dot{q}_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \ddot{q}_i \\ &= \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) \cdot \dot{q}_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \frac{d}{dt}(\dot{q}_i) \\ &= \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{q}_i\right),\end{aligned}$$

where we have applied $\frac{\partial \mathcal{L}}{\partial t} = 0$ as the Lagrangian does not explicitly depend on t. Shifting $\frac{d\mathcal{L}}{dt}$ to the right-hand side,

$$\begin{aligned}\frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{q}_i - \mathcal{L}\right) &= 0 \\ \implies \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{q}_i - \mathcal{L} &= H\end{aligned}$$

for some constant H. How does this obscure-looking term reduce to the familiar $H = T + V$? The answer: Euler's theorem of homogeneous functions.

Euler's Theorem of Homogeneous Functions

Let $f(q_1, q_2, \ldots, q_n)$ be a homogeneous function of degree k such that

$$f(mq_1, mq_2, \ldots, mq_n) = m^k f(q_1, q_2, \ldots, q_n).$$

We claim that

$$\sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \cdot q_i = kf.$$

We differentiate the first equation with respect to m to obtain

$$\begin{aligned} km^{k-1}f(q_1, q_2, \ldots, q_n) &= \frac{\partial f(mq_1, mq_2, \ldots, mq_n)}{\partial m} \\ &= \sum_{i=1}^{n} \frac{\partial f(mq_1, mq_2, \ldots, mq_n)}{\partial (mq_i)} \cdot \frac{\partial (mq_i)}{\partial m} \\ &= \sum_{i=1}^{n} \frac{\partial f(mq_1, mq_2, \ldots, mq_n)}{\partial (mq_i)} \cdot q_i. \end{aligned}$$

Substituting $m = 1$,

$$\sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \cdot q_i = kf.$$

Returning to our problem at hand, we need to evaluate

$$\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{q}_i.$$

Assuming that the potential energy V does not depend on the time-derivative of generalized coordinates — an assumption which is only invalid in the presence of charges, the Lagrangian of a closed system is of the form

$$\mathcal{L} = T(\boldsymbol{q}, \dot{\boldsymbol{q}}) - V(\boldsymbol{q}),$$

where T is quadratic in $\dot{\boldsymbol{q}}$ (i.e. it is a homogeneous function in $\dot{\boldsymbol{q}}$ of degree 2). Therefore,

$$\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{q}_i = \sum_{i=1}^{n} \frac{\partial T}{\partial \dot{q}_i} \cdot \dot{q}_i = 2T$$

$$H = 2T - (T - V) = T + V.$$

We have recovered the familiar expression for mechanical energy! Two assumptions were made in our derivation. Firstly, the Lagrangian is time-independent. This is equivalent to saying that a translation in time does not modify the system's behaviour. Secondly, the potential energy V is independent of $\dot{\boldsymbol{q}}$. When an electromagnetic field is present, this assumption is invalid and energy is seemingly not conserved.

12.5.2 *Cyclic Coordinates*

Observe that if the Lagrangian does not depend on a certain coordinate q_k (but possibly $\dot{q}_k$), it follows from the E-L equations that

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k}\right) = 0$$
$$\implies \frac{\partial \mathcal{L}}{\partial \dot{q}_k} = c$$

is a conserved quantity. Such coordinates q_k are known as **cyclic coordinates**. To illustrate this, consider a particle under the influence of a central potential in spherical coordinates. Then,

$$\mathcal{L} = \frac{1}{2}m\left(\dot{r}^2 + r^2\sin^2\theta\dot{\phi}^2 + r^2\dot{\theta}^2\right) - U(r).$$

Noticing that $\mathcal{L}$ does not depend on ϕ,

$$\frac{\partial \mathcal{L}}{\partial \dot{\phi}} = mr^2\sin^2\theta\dot{\phi}$$

is conserved. This is simply the z-component of the angular momentum of the particle!

12.5.3 *Continuous Symmetries*

Even in the case where there are no obvious cyclic coordinates, there can still be certain conserved quantities. Well, perhaps we just did not choose a convenient coordinate system. Let us consider an instructive Lagrangian of two particles with coordinates q_1 and q_2 that move only in a single direction.

$$\mathcal{L} = \frac{1}{2}m(\dot{q}_1^2 + \dot{q}_2^2) + U(aq_1 + bq_2)$$

for some constants a and b. Observe that if we increment q_1 by a certain $b\delta$ where δ is a constant infinitesimal quantity and q_2 by $-a\delta$,

$$q_1 \to q_1 + b\delta,$$
$$q_2 \to q_2 - a\delta.$$

The Lagrangian remains the same! This is a form of continuous symmetry (continuous because the Lagrangian is invariant after **infinitesimal** variations of coordinates). Now, we turn to another topic: finding a quantity that is conserved in this system. Consider the canonical momenta p_1 and p_2 conjugate to q_1 and q_2 respectively. From the E-L equations,

$$\frac{dp_1}{dt} = \frac{\partial U(aq_1 + bq_2)}{\partial(aq_1 + bq_2)} \cdot \frac{\partial(aq_1 + bq_2)}{\partial q_1} = aU'(aq_1 + bq_2),$$

$$\frac{dp_2}{dt} = \frac{\partial U(aq_1 + b_q)}{\partial(aq_1 + bq_2)} \cdot \frac{\partial(aq_1 + bq_2)}{\partial q_2} = bU'(aq_1 + bq_2).$$

Notice that

$$b\frac{dp_1}{dt} - a\frac{dp_2}{dt} = 0$$

$$\implies bp_1 - ap_2 = c$$

is a conserved quantity of the system, also known as an **integral of motion**, even though there were no cyclic coordinates! Furthermore, we notice that the coefficients in front of p_1 and p_2 are strangely identical to that of δ when varying q_1 and q_2! We shall now explore the deeper reason behind this coincidence.

12.5.4 *Noether's Theorem*

Noether's theorem elegantly connects the symmetries of a system to conserved quantities through Hamilton's principle, which acts as a mediator. Consider a variation of each coordinate q_i in the following manner:

$$q_i \to q_i + f_i(\boldsymbol{q}, t)\delta$$

$$\dot{q}_i \to \dot{q}_i + \dot{f}_i(\boldsymbol{q}, \dot{\boldsymbol{q}}, t)\delta,$$

where $f_i(\boldsymbol{q}, t)$ is any arbitrary function in terms of the generalized coordinates $\boldsymbol{q}$ and t for each coordinate q_i and δ is a constant infinitesimal quantity. If the Lagrangian is preserved to the first order in δ after such an infinitesimal variation of coordinates,

$$0 = \frac{\partial \mathcal{L}}{\partial \delta} = \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial q_i} \cdot \frac{\partial q_i}{\partial \delta} + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \frac{\partial \dot{q}_i}{\partial \delta}$$

$$= \sum_{i=1}^{n} \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i}\right) \cdot f_i + \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot \dot{f}_i$$

$$= \frac{d}{dt}\left(\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot f_i\right)$$

$$\implies \sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot f_i = Q$$

is an integral of motion. Note that we have applied the E-L equation in obtaining the second equality. This relationship is known as Noether's theorem. We see that it elegantly implies that symmetries are closely tied to conserved quantities and this relationship is enforced by Hamilton's principle. In practice, we have to determine f_i by trial and error, but this is usually not too tedious.

Consider the Lagrangian of a two-dimensional harmonic oscillator

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) - \frac{1}{2}k(x^2 + y^2).$$

Let us examine the ramifications of the following variation:

$$x \to x - y\delta,$$

$$y \to y + x\delta.$$

Notice that

$$(x - y\delta)^2 + (y + x\delta)^2 = x^2 + y^2,$$

where we have neglected second-order terms in δ. Meanwhile, $\dot{x}^2 + \dot{y}^2$ remains unchanged as the variations are time-independent. Therefore, the above variation is a symmetry of the system. The conserved quantity is then

$$\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \cdot f_i = \frac{\partial \mathcal{L}}{\partial \dot{x}} \cdot -y + \frac{\partial \mathcal{L}}{\partial \dot{y}} \cdot x = m(x\dot{y} - y\dot{x}).$$

This is, again, simply the angular momentum of the oscillating body in the z-direction. We have merely expressed our system in the cumbersome Cartesian coordinates instead of the convenient polar coordinates, which would have directly resulted in a cyclic angular coordinate!

Conservation of Momentum

Now, let us derive two conserved quantities in Newtonian mechanics from the homogeneity and isotropy of space and time in inertial frames. The conservation of momentum is a result of the homogeneity of space — which simply put, means that space is uniform. This causes the Lagrangian to be invariant when the entire closed system is displaced by a certain distance in a certain direction. Let there be n particles in total, with the ith particle having coordinates x_i, y_i and z_i. Then, we can shift the entire system (or our coordinate axes) such that

$$x_i \to x_i + \delta,$$

$$y_i \to y_i,$$

$$z_i \to z_i,$$

and the Lagrangian should still remain the same. By Noether's theorem, the quantity

$$\sum_{i=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{x}_i} = \sum_{i=1}^{n} \frac{\partial T}{\partial \dot{x}_i} = \sum_{i=1}^{n} m\dot{x}_i = p_x,$$

which corresponds to the x-component of the total momentum of the system, is conserved for a closed system. Note that in writing the first equality, the potential V was assumed to be independent of $\dot{x}_i$. We can repeat the above process for the y and z directions, to conclude that the total momentum vector $\boldsymbol{p}$ of a closed system is conserved.

Conservation of Angular Momentum

Due to the isotropy of space in an inertial frame (that is, all directions are equal), the Lagrangian of a closed system does not change after an infinitesimal rotation about an origin. Note that the position vector of the ith particle, $\boldsymbol{r}_i$, changes in the following manner after an infinitesimal rotation about the origin described by the rotation vector $\boldsymbol{\delta}$.

$$\boldsymbol{r}_i \to \boldsymbol{r}_i + \boldsymbol{\delta} \times \boldsymbol{r}_i.$$

Assuming that the rotation is solely along the z-direction, $\boldsymbol{\delta} = \delta\hat{\boldsymbol{z}}$,

$$x_i \to x_i - y_i\delta,$$

$$y_i \to y_i + x_i\delta,$$

$$z_i \to z_i.$$

The conserved quantity is then

$$\begin{aligned}\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{x}_i}\cdot(-y_i)+\sum_{i=1}^{n}\frac{\partial\mathcal{L}}{\partial\dot{y}_i}\cdot x_i &= \sum_{i=1}^{n} m(x_i\dot{y}_i-y_i\dot{x}_i)\\ &= \sum_{i=1}^{n}(\boldsymbol{r}_i\times\boldsymbol{p}_i)_z\\ &= L_z,\end{aligned}$$

which is the z-component of the total angular momentum of the system. We can then repeat this for rotations about the x and y-axes to conclude that the total angular momentum $\boldsymbol{L}$ is conserved for a closed system.

Problems

1. *Pendulum about Rotating Pivot**

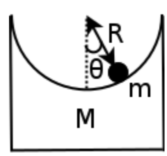
A pendulum bob of mass m is initially attached to the circumference of a wheel of radius R at polar coordinates $(R, 0)$ via an inextensible string of length l. The x and y directions are defined to be positive rightwards and upwards respectively. The wheel then begins to rotate anti-clockwise at a constant angular velocity ω. Determine the equations of motion of the bob.

2. *Cylinder on Inclined Plane***

A uniform cylinder of mass m and radius R is initially held still on an inclined plane with an angle of inclination θ and mass M, with its cylindrical axis parallel to the width of the plane. If the cylinder is released and subsequently rolls without slipping on the surface of the plane, determine the horizontal acceleration of the inclined plane.

3. *Small Oscillations on Circle***

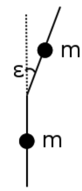
A circular track of radius R is carved from a rectangular block, as shown in the figure below. The final track has mass M. Define θ to be the angular coordinate of a mass m that is placed on the frictionless track. For small θ and small velocities, determine the angular frequency of small oscillations of m about its equilibrium position.



Now, consider a separate problem where a bead of mass m is hung along a thin circular hoop of mass M and radius R that rests vertically on the ground. There is no friction between m and the hoop. If the mass is released at a small angle from the bottom of the hoop and the hoop subsequently rolls without slipping on the ground, determine the angular frequency of small oscillations about the bottom without performing additional calculations.

4. *Two Sticks***

Two massless sticks of length $2r$, each with a mass m fixed at its center, are connected via a hinge at its ends. Referring to the figure on the next page, the bottom end of the lower stick is hinged to the ground. Initially, the sticks are held such that the lower stick is vertical while the upper one is tilted at a small angle ε clockwise from the vertical. At the moment where they are subsequently released, determine the instantaneous angular accelerations of the two sticks.

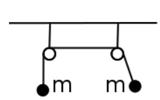


5. *Particle on Rotating Plate***

A particle of mass m rests on a frictionless plate that is initially horizontal. Then, you lift up the left edge of the plate and rotate it at a constant clockwise angular velocity ω about its fixed right edge. Denote the distance between m and the right edge as x. Determine $x(t)$ given that $x = x_0$ at $t = 0$.

6. *Swinging Pulley***

Two equal masses m are connected by a string that hangs over two fixed pulleys (of negligible size) as shown in the figure below. The left mass moves along the vertical direction while the right one is able to swing back and forth in the plane of the masses and pulleys. Find the equations of motion of this system. Assuming that the left mass starts at rest and the right mass undergoes small oscillations with angular amplitude $\varepsilon \ll 1$, what is the initial average acceleration (averaged over a few periods of the mass on the right) of the mass on the left? In which direction is this acceleration oriented towards?



7. *Particle on Rotating Hoop***

A particle is constrained to move on a vertical thin hoop of radius R that initially lies in the xz-plane and is rotating about the vertical z-axis at a constant angular velocity ω anti-clockwise. Define the origin to be at the center of the hoop. Let θ be the angular coordinate of the particle, measured clockwise from the positive z-axis. Show that there exist angular positions, other than that at the top and bottom of the hoop, such that the particle can remain stationary if the magnitude of ω is large enough. Find these angular positions and the minimum value of ω required. Do these equilibrium positions correspond to stable equilibria? If so, determine the frequency of small angular oscillations if the particle is slightly displaced from those angular positions.

8. *Particle on Rotating Curve***

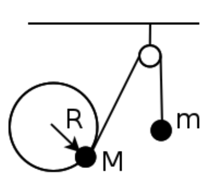
A particle is constrained to move on a rigid curve $f(x)$ that is initially in the $x \geq 0$ region in the x-z plane. The curve rotates at a constant angular velocity ω about the z-axis in the anti-clockwise direction. Obtain an equation that can be used to determine possible initial x-coordinates that correspond to equilibrium positions. Determine the condition for such positions to be stable equilibria.

9. *Massive Spring Oscillation***

Firstly, prove that if a spring of mass M undergoes uniform stretching, the kinetic energy of the spring is $\frac{1}{6}M\dot{\varepsilon}^2$, where ε is the extension of the spring. Now, determine the frequency of oscillations for a massive spring of mass M and spring constant k, connected to a point mass m and a wall, on a horizontal table. Lastly, suppose that we add an identical spring to the system. How does the frequency of oscillations change in each case if the additional spring is connected in series and in parallel with respect to the original one?

10. *Oscillations with Hoop***

A particle of mass M is attached to a massless hoop of radius R as shown in the figure below. The hoop is supported about a fixed axle through its center but it is able to rotate freely. M is then connected to another particle of mass m via a massless string hung over a frictionless, small pulley. Determine the angular frequency of small oscillations of this system about its equilibrium state if m always remains vertical.



11. *Oscillating Mass Attached to Another Mass***

A simple pendulum of mass m and length l is attached to a particle of mass M that lies on a pair of frictionless horizontal rails that extend in a single direction. Find the equations of motion of this system if m only oscillates in the plane of the rails. For small oscillations of m about the vertical, determine the normal modes of this motion and explain them.

12. *Atwood's Machine with Massive String***

Consider the simplest Atwood's Machine with two masses m_1 and m_2, connected by a uniform inextensible string of mass m_3 and length l. Define the x-axis to be positive downwards and let the coordinate of m_1 be x. If the circumference of the pulley is assumed to be negligible such that the amount of string wrapping around the pulley is negligible, find the constant $\alpha > 0$, if $x(t)$ can be expressed as

$$x(t) = Ae^{\alpha t} + Be^{-\alpha t} + C,$$

for some constants A, B and C.

13. *Brachistochrone***

In the xy-plane, a bead starts off from rest at the origin and slides along a smooth wire to a predetermined point below the origin. Show that the shape of the wire $y(x)$ that minimizes the time taken by this bead satisfies the parametric equations

$$x = a(\phi - \sin\phi),$$
$$y = a(1 - \cos\phi),$$

where y is defined to be positive downwards, a is a constant and ϕ is a variable. Try to deduce the parameterization yourself! These equations describe the shape of a cycloid which is the trajectory of a particle attached to a rim of a circle of radius a that rolls without slipping in the x-direction. Hint: Apply a certain "conservation law."

14. *Wire Pendulum***

A particle of mass m is constrained to move along a wire described by the equation $y(x)$, which is concave upwards. The wire is horizontal at the origin. Let $s(x)$ denote the arc length between the point on the wire at x-coordinate x and the origin along the wire, such that the particle's velocity along the wire is $\dot{s}$. Show that if $y = ks^2$ for some constant k, the particle will undergo simple harmonic motion, regardless of its initial distance from the origin and velocity. Solve for y and x in terms of parametric equations and describe the shape of the wire. We have thus constructed a pendulum that exhibits simple harmonic motion exactly, independent of the amplitude of oscillation.

15. *Virial Theorem**

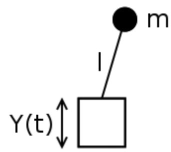
Consider an arbitrary bounded system of N particles (bounded implies that the coordinates and velocities of these particles do not diverge) subject to a homogeneous potential energy $U(x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_N, y_N, z_N)$ of degree d where x_i, y_i and z_i are the x, y and z-coordinates of the ith particle. Consider the quantity $G = \sum_{i=1}^{N} m_i \boldsymbol{r}_i \cdot \dot{\boldsymbol{r}}_i$ where m_i and $\boldsymbol{r}_i$ are the mass and position vector of the ith particle. Argue why $\langle \frac{dG}{dt} \rangle = \lim_{t \to \infty} \frac{1}{t} \int_0^t \frac{dG}{dt} dt = 0$ where angle brackets in this problem mean that the time average is taken over a period that tends to infinity. Hence, prove that $\langle 2T - dU \rangle = 0$ where T is the total kinetic energy of the system. You may find Euler's theorem of homogeneous functions to be very useful.

16. *Cylinder on Cylinder***

Consider a uniform cylinder of mass m and radius r_2 placed on top of another uniform cylinder of mass M and radius r_1 that is fixed translationally but can still rotate about its cylindrical axis. Both of the cylindrical axes are aligned and are parallel to the ground. The cylinder on top is then given a slight azimuthal push. Let the angle subtended by the line joining the two centers of the cylinder and the vertical be θ (positive clockwise). Using the method of Lagrange multipliers, find $\ddot{\theta}$ in terms of θ and find θ_{max}, the angle at which the cylinders lose contact with each other. Assume that the cylinders do not slip with respect to each other.

17. *Inverted Pendulum***

A pendulum consists of a mass m attached to a platform oscillating vertically at $Y(t) = A \cos \omega t$ via a massless stick of length l. Determine the equation of motion of m, in terms of the angle θ that the stick subtends with the vertical. Surprisingly, when $A \ll l$ and ω is large enough, the stick will not topple when m is given a slight angular deviation from the vertical when it is initially on top of the platform (i.e. the pendulum is upside-down). Instead, it will sort of oscillate like an upside-down pendulum. Make rough arguments about why this should be the case. Your calculations do not have to be exact. Hint: as ω is large, θ oscillates rapidly with angular frequency ω over a short period (due to the movement of the platform) such that we can write $\theta = B + C\cos\omega t$ where B and C are possibly time-dependent constants whose time-scales are much larger than $\frac{2\pi}{\omega}$. Furthermore, we know that $C \ll B$ as $A \ll l$ such that the oscillation of the platform should not affect the position of m much in a single period.



18. *Double Pendulum***

A pendulum bob of mass m_1 is fixed via an inextensible string of length l_1 to a ceiling. Then, another bob of mass m_2 is tied to this first bob via another string of length l_2. Let the anti-clockwise angles subtended by the first and second strings and the vertical be θ_1 and θ_2 respectively (the strings and the vertical lie in a single plane). Find the possible frequencies of small oscillations.

19. *Unraveling Rope***

A uniform, smooth rope of total length l and mass M is initially wrapped into a stationary spiral of radius R. The exterior end, which is at the bottom of the spiral, is fixed to a support as shown in the figure below. The rope is then given a slight push such that it begins to roll without slipping on the ground in a single direction, while unraveling as it travels. Assume that the spiral constantly maintains a circular shape and that segments of the rope that are no longer part of the spiral immediately come to a stop. Write down the Lagrangian of this system and derive the equation of motion. Then, show that given the slightest impulsive push, the rope will always completely unravel. Finally, determine the time taken for it to completely unravel if the initial velocity of the center of mass of the rope is negligible. As the result will involve a difficult integral, use the constant α to represent $\int_0^{\frac{\pi}{2}} \sqrt{\sin\theta}d\theta \approx 1.198$ in your answer.

Solutions

1. Pendulum about Rotating Pivot*

Let the angle that the string makes with the vertical be θ, taking positive values in the anti-clockwise direction. If we define the origin to be at the center of the wheel, the coordinates of the bob are

$$x = R\cos\omega t + l\sin\theta,$$
$$y = R\sin\omega t - l\cos\theta,$$

which implies that

$$\dot{x} = l\cos\theta\dot{\theta} - R\omega\sin\omega t,$$
$$\dot{y} = l\sin\theta\dot{\theta} + R\omega\cos\omega t.$$

Hence, the squared speed of the bob is

$$v^2 = \dot{x}^2 + \dot{y}^2 = l^2\dot{\theta}^2 + 2Rl\omega\sin(\theta - \omega t)\dot{\theta} + R^2\omega^2.$$

Hence, the Lagrangian of the bob (after discarding the $\frac{1}{2}mR^2\omega^2$ and $mgR\sin\omega t$ terms which represent total time derivatives) is

$$\mathcal{L} = \frac{1}{2}m\left(l^2\dot{\theta}^2 + 2Rl\omega\sin(\theta - \omega t)\dot{\theta}\right) + mgl\cos\theta.$$

Applying the E-L equation with respect to θ,

$$ml^2\ddot{\theta} - mRl\omega^2\cos(\theta - \omega t) + mRl\omega\cos(\theta - \omega t)\dot{\theta} = -mgl\sin\theta,$$

which returns to the familiar $ml^2\ddot{\theta} = -mgl\sin\theta$ in the limiting case where $\omega = 0$.

2. Cylinder on Inclined Plane**

Let the horizontal coordinate of the vertical edge of the plane be x. Let s be the distance between the point of contact of the cylinder with the hypotenuse and the top of the vertical edge. Then, the coordinates of the center of the cylinder are $(x + s\cos\theta + R\sin\theta, -s\sin\theta + R\cos\theta)$ if we presume the slope of the plane to be negative in the positive x-direction. The translational velocities of the center of mass of the cylinder in the x and y directions are $\dot{x} + \dot{s}\cos\theta$ and $-\dot{s}\sin\theta$ respectively. Since the cylinder rolls without slipping, $R\omega = \dot{s}$ where ω is the angular velocity of the cylinder. The rotational kinetic energy of the cylinder is $\frac{1}{2}I\omega^2 = \frac{1}{4}mR^2\omega^2 = \frac{1}{4}m\dot{s}^2$. Hence, the Lagrangian of the system comprising the cylinder and plane is

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m\left((\dot{x}+\dot{s}\cos\theta)^2 + \dot{s}^2\sin^2\theta\right) + \frac{1}{2}M\dot{x}^2 + \frac{1}{4}m\dot{s}^2 + mgs\sin\theta \\ &= \frac{1}{2}m\left(\dot{x}^2+\dot{s}^2+2\dot{x}\dot{s}\cos\theta\right) + \frac{1}{2}M\dot{x}^2 + \frac{1}{4}m\dot{s}^2 + mgs\sin\theta,\end{aligned}$$

where we have discarded the total time derivative $-mgR\cos\theta$. The E-L equations with respect to x and s yield

$$m\ddot{x} + m\ddot{s}\cos\theta + M\ddot{x} = 0,$$
$$\frac{3}{2}m\ddot{s} + m\ddot{x}\cos\theta = mg\sin\theta.$$

Solving this system of equations,

$$\ddot{x} = -\frac{mg\sin\theta\cos\theta}{\left(\frac{1}{2}+\sin^2\theta\right)m + \frac{3}{2}M}.$$

3. Small Oscillations on Circle**

Define the origin at the center of the circle and the x and y-axes to be positive rightwards and upwards respectively. Then, define x as the x-coordinate of the center of mass of the track. The coordinates of the mass m are then $(x + R\sin\theta, -R\cos\theta)$. The squared speed of the mass m is

$$v^2 = (\dot{x} + R\cos\theta\dot{\theta})^2 + R^2\sin^2\theta\dot{\theta}^2 = \dot{x}^2 + R^2\dot{\theta}^2 + 2R\cos\theta\dot{x}\dot{\theta}.$$

The Lagrangian of the combined system comprising the track and the mass is

$$\mathcal{L} = \frac{1}{2}m(\dot{x}^2 + R^2\dot{\theta}^2 + 2R\cos\theta\dot{x}\dot{\theta}) + \frac{1}{2}M\dot{x}^2 + mgR\cos\theta.$$

Applying the E-L equation with respect to x and θ,

$$m\ddot{x} - mR\sin\theta\dot{\theta}^2 + mR\cos\theta\ddot{\theta} + M\ddot{x} = 0,$$
$$-mR\sin\theta\dot{x}\dot{\theta} + mR\cos\theta\ddot{x} + mR^2\ddot{\theta} = -mR\sin\theta\dot{x}\dot{\theta} - mgR\sin\theta,$$
$$\implies mR\ddot{\theta} = -mg\sin\theta - m\cos\theta\ddot{x}.$$

For small angles, $\sin\theta \approx \theta$ and $\cos\theta \approx 1$. Furthermore, discarding the $\dot{\theta}^2$ term which is very small for small angles by the conservation of energy,

$$(m+M)\ddot{x} = -mR\ddot{\theta},$$
$$mR\ddot{\theta} = -mg\theta - m\ddot{x}.$$

Eliminating $\ddot{x}$,

$$\ddot{\theta} = -\frac{g(m+M)}{RM}\theta.$$

The angular frequency of oscillations is thus

$$\omega = \sqrt{\frac{g(m+M)}{RM}}.$$

In the second problem, define x to be the horizontal coordinate of the center of mass of the circular hoop and θ to be the anti-clockwise angular coordinate of the small mass from the bottom of the hoop. Then, the Lagrangian of this system is almost the same as the previous one, with the addition of a rotational kinetic energy term of the hoop. The moment of inertia of a hoop is MR^2. Therefore, its rotational kinetic energy with angular velocity ω is

$$\frac{1}{2}MR^2\omega^2 = \frac{1}{2}M\dot{x}^2,$$

after applying the non-slip condition $R\omega = \dot{x}$. This implies that the kinetic energy of M doubles, as compared to the previous question. Since the only dependence of the Lagrangian on M is through the kinetic energy term, the angular frequency Ω in the second problem can be directly determined by substituting $2M$ for M in the result, ω, of the first problem, and

$$\Omega = \sqrt{\frac{g(m+2M)}{2RM}}.$$

4. Two Sticks**

Label the bottom and top masses as 1 and 2 respectively. Define the x and y-axes to be positive rightwards and upwards, with the origin located at the bottom end of the lower stick. Next, define θ_1 and θ_2 as the clockwise angles that the lower and upper sticks subtend with the vertical. The coordinates of masses 1 and 2 are

$$\begin{aligned}
&(x_1, y_1) = (r\sin\theta_1, r\cos\theta_1)\\
\implies &(\dot{x}_1, \dot{y}_1) = (r\cos\theta_1\dot{\theta}_1, -r\sin\theta_1\dot{\theta}_1),\\
&(x_2, y_2) = (2r\sin\theta_1 + r\sin\theta_2, 2r\cos\theta_1 + r\cos\theta_2)\\
\implies &(\dot{x}_2, \dot{y}_2) = (2r\cos\theta_1\dot{\theta}_1 + r\cos\theta_2\dot{\theta}_2, -2r\sin\theta_1\dot{\theta}_1 - r\sin\theta_2\dot{\theta}_2).
\end{aligned}$$

The Lagrangian of the entire system is thus

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m(\dot{x}_1^2+\dot{y}_1^2)+\frac{1}{2}m(\dot{x}_2^2+\dot{y}_2^2)-mgr\cos\theta_1-mg(2r\cos\theta_1+r\cos\theta_2)\\ &= \frac{1}{2}mr^2\dot{\theta}_1^2+\frac{1}{2}m\left[4r^2\dot{\theta}_1^2+r^2\dot{\theta}_2^2+4r^2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2)\right]\\ &\quad -mgr(3\cos\theta_1+\cos\theta_2).\end{aligned}$$

Applying the E-L equation with respect to θ_1 and θ_2, while substituting $\dot{\theta}_1=\dot{\theta}_2=0$ initially, we obtain

$$3mgr\sin\theta_1 = 5mr^2\ddot{\theta}_1+2mr^2\ddot{\theta}_2\cos(\theta_1-\theta_2),$$
$$mgr\sin\theta_2 = mr^2\ddot{\theta}_2+2mr^2\ddot{\theta}_1\cos(\theta_1-\theta_2).$$

Substituting $\theta_1=0$ and $\theta_2=\varepsilon\ll 1$ initially,

$$5\ddot{\theta}_1+2\ddot{\theta}_2=0,$$
$$2\ddot{\theta}_1+\ddot{\theta}_2=\frac{g\varepsilon}{r}.$$

Solving the above simultaneously, the instantaneous angular accelerations are initially

$$\ddot{\theta}_1=-\frac{2g\varepsilon}{r},$$
$$\ddot{\theta}_2=\frac{5g\varepsilon}{r}.$$

5. Particle on Rotating Plate**

Define the origin at the fixed right edge of the plate and consider polar coordinates about this origin. x effectively functions as the radial coordinate of the particle m. Therefore, it has a radial velocity $\dot{x}$ and tangential velocity $x\omega$. The Lagrangian of m is thus

$$\mathcal{L}=\frac{1}{2}m(\dot{x}^2+x^2\omega^2)-mgx\sin\omega t.$$

Applying the E-L equation,

$$m\ddot{x}=mx\omega^2-mg\sin\omega t$$
$$\ddot{x}-\omega^2x=-g\sin\omega t.$$

We first solve for the particular solution of this expression. Substituting a trial solution of the form $x=A\sin(\omega t+\phi)$,

$$-2A\omega^2\sin(\omega t+\phi)=-g\sin\omega t.$$

Considering the instant where $t = 0$, we obtain

$$\phi = 0.$$

Equating the amplitudes of both sides,

$$A = \frac{g}{2\omega^2}.$$

Therefore, the particular solution is

$$x_p = \frac{g}{2\omega^2}\sin\omega t.$$

Moving on, we investigate the homogeneous solution to the equation

$$\ddot{x} - \omega^2 x = 0,$$

whose solution is

$$x_h = Be^{\omega t} + Ce^{-\omega t}$$

for some constants B and C. The general solution for $x(t)$ is thus

$$x = x_h + x_p = Be^{\omega t} + Ce^{-\omega t} + \frac{g}{2\omega^2}\sin\omega t.$$

Imposing the initial conditions $B + C = x_0$ and $\omega B - \omega C + \frac{g}{2\omega} = 0$,

$$B = \frac{x_0}{2} - \frac{g}{4\omega^2},$$
$$C = \frac{x_0}{2} + \frac{g}{4\omega^2},$$
$$x = \left(\frac{x_0}{2} - \frac{g}{4\omega^2}\right)e^{\omega t} + \left(\frac{x_0}{2} + \frac{g}{4\omega^2}\right)e^{-\omega t} + \frac{g}{2\omega^2}\sin\omega t.$$

6. Swinging Pulley**

Define the instantaneous length of the string segment between the right pulley and the right mass to be r. Then, the vertical coordinate (positive upwards) of the left mass is $-r$ up to the addition of a constant (whose associated gravitational potential energy vanishes from the Lagrangian anyway). Let θ be the anti-clockwise angle subtended by the right mass and the vertical. The Lagrangian of this system is

$$\mathcal{L} = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2) + mgr(\cos\theta - 1).$$

Applying the E-L equation with respect to r and θ,

$$2m\ddot{r} = mr\dot{\theta}^2 + mg(\cos\theta - 1)$$
$$2m\dot{r}\dot{\theta} + mr^2\ddot{\theta} = -mgr\sin\theta.$$

When the right mass undergoes small oscillations, θ, $\dot{\theta}$ and $\dot{r}$ are small (initially) such that the second equation becomes

$$\ddot{\theta} = -\frac{g}{r}\theta.$$

If $\dot{r}$ is small, r is approximately constant over a period of θ. Therefore,

$$\theta = \varepsilon\cos\left(\sqrt{\frac{g}{r}}t + \phi\right)$$
$$\implies \dot{\theta} = -\varepsilon\sqrt{\frac{g}{r}}\sin\left(\sqrt{\frac{g}{r}}t + \phi\right),$$

where ε is the amplitude and ϕ is a constant. Substituting these into the first equation of motion while using the Maclaurin expansion $\cos\theta \approx 1 - \frac{\theta^2}{2}$ for small θ,

$$2m\ddot{r} = mg\varepsilon^2\sin^2\left(\sqrt{\frac{g}{r}}t + \phi\right) - \frac{1}{2}mg\dot{\theta}^2$$
$$\ddot{r} = \frac{g\varepsilon^2}{2}\sin^2\left(\sqrt{\frac{g}{r}}t + \phi\right) - \frac{g\varepsilon^2}{4}\cos^2\left(\sqrt{\frac{g}{r}}t + \phi\right).$$

Since the averages of both $\sin^2(\sqrt{\frac{g}{r}}t + \phi)$ and $\cos^2(\sqrt{\frac{g}{r}}t + \phi)$ are $\frac{1}{2}$ over a single period, the initial average $\ddot{r}$ is

$$\langle\ddot{r}\rangle = \frac{g\varepsilon^2}{8}.$$

The positive value of $\langle\ddot{r}\rangle$ indicates that the left mass is accelerated upwards.

7. Particle on Rotating Hoop**

The velocity of the particle along the hoop is $R\dot{\theta}$ while the velocity of the particle perpendicular to the plane of the hoop is $R\sin\theta\omega$ (due to its rotation).

Hence, the Lagrangian of the particle is

$$\mathcal{L} = \frac{1}{2}m\left(R^2\dot{\theta}^2 + R^2\sin^2\theta\omega^2\right) - mgR\cos\theta.$$

From the E-L equation,

$$mR^2\ddot{\theta} = mR^2\sin\theta\cos\theta\omega^2 + mgR\sin\theta. \tag{12.10}$$

At the equilibrium positions, $\ddot{\theta} = 0$. Hence,

$$\cos\theta = -\frac{g}{R\omega^2}$$

if $\sin\theta \neq 0$ (excluding positions corresponding to the top and bottom of the hoop). Since $|\cos\theta| \leq 1$,

$$\omega^2 \geq \frac{g}{R}$$
$$|\omega| \geq \sqrt{\frac{g}{R}}$$

for such an equilibrium to exist. Then, the possible equilibrium angles satisfy

$$\cos\theta = -\frac{g}{R\omega^2}.$$

The two equilibrium positions correspond to θ_0 and $2\pi - \theta_0$ with

$$\theta_0 = \cos^{-1}\left(-\frac{g}{R\omega^2}\right).$$

We shall just analyze the stability of θ_0 as the set-up is symmetrical about the z-axis. In this process, remember that from Eq. (12.10),

$$\frac{1}{2}mR^2\sin 2\theta_0\omega^2 + mgR\sin\theta_0 = 0,$$

as θ_0 corresponds to an equilibrium position. Suppose we displace the particle slightly such that its angular coordinate $\theta = \theta_0 + \varepsilon$ for some infinitesimal angular displacement ε. Substituting this expression into the equation of motion,

$$mR^2\ddot{\varepsilon} = \frac{1}{2}mR^2\sin(2\theta_0 + 2\varepsilon)\omega^2 + mgR\sin(\theta_0 + \varepsilon).$$

Using small angle approximations $\sin x \approx x$ and $\cos x \approx 1$, valid if x is small, and using the trigonometric identity $\sin(a+b) = \sin a\cos b + \sin b\cos a$,

$$mR^2\ddot{\varepsilon} = \frac{1}{2}mR^2\sin 2\theta_0\omega^2 + mR^2\omega^2\cos 2\theta_0\varepsilon + mgR\sin\theta_0 + mgR\cos\theta_0\varepsilon.$$

Note that

$$\cos 2\theta_0 = 2\cos^2\theta_0 - 1 = \frac{2g^2}{R^2\omega^4} - 1.$$

Substituting the expressions for $\cos\theta_0$ and $\cos 2\theta_0$ and canceling the terms $\frac{1}{2}mR^2\sin 2\theta_0\omega^2 + mgR\sin\theta_0 = 0$,

$$\ddot{\varepsilon} = \left(\frac{g^2}{R^2\omega^2} - \omega^2\right)\varepsilon = \frac{1}{\omega^2}\left(\frac{g^2}{R^2} - \omega^4\right)\varepsilon.$$

Note that the coefficient in front of ε is negative as $\omega^2 > \frac{g}{R}$. This means that this equilibrium corresponds to a stable equilibrium. The frequency of small oscillations about this position is then

$$\Omega = \frac{\sqrt{R^2\omega^4 - g^2}}{R|\omega|}.$$

8. Particle on Rotating Curve**

Let x be the horizontal coordinate of the particle in the plane of the curve. The component of the velocity of the particle along the curve is $\sqrt{1 + f'(x)^2}\dot{x}$ while the component perpendicular to the plane of the curve is $x\omega$. Hence, the Lagrangian of the particle is

$$\mathcal{L} = \frac{1}{2}m\left[\left(1 + f'(x)^2\right)\dot{x}^2 + x^2\omega^2\right] - mgf(x).$$

The E-L equation yields

$$m(1 + f'(x)^2)\ddot{x} + 2mf'(x)f''(x)\dot{x}^2 = mx\omega^2 - mgf'(x). \tag{12.11}$$

At equilibrium positions, $\ddot{x} = 0$ and $\dot{x} = 0$ and we get

$$\frac{x}{f'(x)} = \frac{g}{\omega^2}.$$

Let the x-coordinate of a particular equilibrium position be x_0. Then, to determine whether this position corresponds to a stable equilibrium, consider an infinitesimal displacement ε such that the x-coordinate of the particle becomes $x_0 + \varepsilon$. Substituting this expression into Eq. (12.11),

$$\{1 + [f'(x_0) + f''(x_0)\varepsilon]^2\}\ddot{\varepsilon} + 2[f'(x_0) + f''(x_0)\varepsilon][f''(x_0) + f'''(x_0)\varepsilon]\dot{\varepsilon}^2$$
$$= (x_0 + \varepsilon)\omega^2 - g[f'(x_0) + \varepsilon f''(x_0)],$$

where we have performed a Taylor expansion and approximated $f'(x_0 + \varepsilon) \approx f'(x_0) + f''(x_0)\varepsilon$ — ditto for subsequent derivatives. Simplifying the resultant equation by cancelling terms which correspond to the equation of motion at coordinate x_0 that equate to zero, and by discarding second-order infinitesimal terms (note that $\dot{\varepsilon}$ and $\ddot{\varepsilon}$ are small too),

$$\ddot{\varepsilon} = \frac{\omega^2 - gf''(x_0)}{1 + f'(x_0)^2}\varepsilon.$$

Hence, the frequency of small oscillations is

$$\Omega = \sqrt{\frac{gf''(x_0) - \omega^2}{1 + f'(x_0)^2}}$$

if $gf''(x_0) > \omega^2$. This inequality is also the condition for x_0 to be a stable equilibrium.

9. Massive Spring Oscillation**

Let the relaxed length of the spring be l and define the origin at the fixed end of the spring. Then, after an extension ε, an infinitesimal segment of the spring which was initially between coordinates x and $x+dx$, now corresponds to coordinates

$$x' = x\left(1 + \frac{\varepsilon}{l}\right)$$

and $x' + dx'$. Hence, the velocity of this segment as a function of $\dot{\varepsilon}$ is

$$\dot{x}' = \frac{x}{l}\dot{\varepsilon}.$$

The kinetic energy of this infinitesimal segment is $\frac{1}{2}\lambda\dot{x}'^2 dx$ where $\lambda = \frac{M}{l}$ is the density of the relaxed spring and dx is the length of the infinitesimal segment when the spring was relaxed. Note that the mass of this segment, after stretching, is still λdx as no mass crosses its boundaries. Hence, the total kinetic energy of the spring is obtained by integrating $\frac{1}{2}\lambda\dot{x}'^2 dx$ over the original length of the entire spring.

$$T = \int_0^l \frac{1}{2}\lambda \dot{x}'^2 dx = \int_0^l \frac{1}{2}\lambda \frac{x^2}{l^2}\dot{\varepsilon}^2 dx = \frac{1}{6}\lambda l \dot{\varepsilon}^2 = \frac{1}{6}M\dot{\varepsilon}^2.$$

Now, consider a horizontal spring-mass system in which the spring can be presumed to stretch uniformly. Define the origin at the stationary end of the spring again. Let x be the horizontal coordinate of the mass. Then, the Lagrangian of the spring-mass system is

$$\mathcal{L} = \frac{1}{2}m\dot{x}^2 + \frac{1}{6}M\dot{x}^2 - \frac{1}{2}k(x-l)^2.$$

The E-L equations yield

$$\left(m + \frac{M}{3}\right)\ddot{x} = -k(x-l)$$
$$\implies \ddot{y} = -\frac{k}{m+\frac{M}{3}}y,$$

where $y = x - l$. The angular frequency of oscillation is

$$\omega = \sqrt{\frac{k}{m+\frac{M}{3}}}.$$

It can be seen that the effect of the mass of the spring is to "add" an additional $\frac{M}{3}$ to the oscillating body. Adding an identical spring doubles the mass of the effective spring. Furthermore, the effective spring constants for two springs connected in series and parallel are $\frac{k}{2}$ and $2k$ respectively. Hence,

$$\omega_{series} = \sqrt{\frac{k}{2m+\frac{4M}{3}}},$$
$$\omega_{parallel} = \sqrt{\frac{2k}{m+\frac{2M}{3}}}.$$

10. Oscillations with Hoop**

Let θ denote the anti-clockwise angle subtended by M and the vertical. Then, the vertical coordinate (taking upwards as positive) of m is $-R\theta$ up to the addition of a constant. Therefore, the Lagrangian of the system is

$$\mathcal{L} = \frac{1}{2}(m+M)R^2\dot{\theta}^2 + mgR\theta + MgR\cos\theta.$$

Applying the E-L equation with respect to θ,

$$(m+M)R^2\ddot{\theta} = gR(m - M\sin\theta).$$

The equilibrium angle θ_0 evidently corresponds to

$$\sin\theta_0 = \frac{m}{M}.$$

Now, consider a slight deviation from this angle such that $\theta = \theta_0 + \varepsilon$ where $\varepsilon \ll 1$. Substituting this expression into the equation of motion,

$$(m+M)R^2\ddot{\varepsilon} = gR(m - M\sin\theta_0 - M\cos\theta_0\varepsilon),$$

where we have applied the trigonometric identity $\sin(x+y) = \sin x\cos y + \cos x\sin y$. Since $m - M\sin\theta_0 = 0$,

$$\ddot{\varepsilon} = -\frac{Mg\cos\theta_0}{(m+M)R}\varepsilon,$$

which describes a simple harmonic motion with angular frequency

$$\omega = \sqrt{\frac{Mg\cos\theta_0}{(m+M)R}} = \sqrt{\frac{Mg\sqrt{1-\frac{m^2}{M^2}}}{(m+M)R}} = \sqrt[4]{\frac{M-m}{M+m}}\sqrt{\frac{g}{R}}.$$

11. Oscillating Mass Attached to Another Mass**

Define the x-axis to be along the rails and let the x-coordinate of M be x. Let θ denote the instantaneous anti-clockwise angle subtended by the pendulum and the vertical. Then, the coordinates of m are $(x + l\sin\theta, -l\cos\theta)$. The Lagrangian of this system is thus

$$\mathcal{L} = \frac{1}{2}M\dot{x}^2 + \frac{1}{2}m(\dot{x}^2 + 2l\cos\theta\dot{x}\dot{\theta} + l^2\dot{\theta}^2) + mgl\cos\theta.$$

Applying the E-L equation with respect to x and θ,

$$(M+m)\ddot{x} - ml\sin\theta\dot{\theta}^2 + ml\cos\theta\ddot{\theta} = 0,$$

$$ml\cos\theta\ddot{x} - ml\sin\theta\dot{x}\dot{\theta} + ml^2\ddot{\theta} + mgl\sin\theta = 0.$$

For small oscillations of m, θ and $\dot{\theta}$ are small such that the above become

$$(M+m)\ddot{x} + ml\ddot{\theta} = 0,$$

$$\ddot{x} + l\ddot{\theta} + g\theta = 0.$$

Eliminating $\ddot{x}$,

$$\ddot{\theta} = -\frac{(M+m)g}{Ml}\theta$$
$$\implies \theta = A\cos(\omega t + \phi)$$

for some constants A and ϕ determined by initial conditions and where

$$\omega = \sqrt{\frac{(M+m)g}{Ml}}$$

is the angular frequency of the simple harmonic motion. Now, integrating $(M+m)\ddot{x} + ml\ddot{\theta} = 0$ twice, we obtain

$$x = -\frac{ml}{M+m}\theta + Bt + C$$
$$x = -\frac{ml}{M+m}A\cos(\omega t + \phi) + Bt + C,$$

where B and C are some constants determined by initial conditions. In the current context, the normal modes refer to the independent motions that can be exhibited by this system (i.e. fix all other motions and see if there is still one more). C is largely irrelevant here as it just depends on the choice of origin. One obvious normal mode occurs when $A = 0$ such that m remains vertical while the two particles simply travel along the rails at a constant velocity. Another mode occurs when $B = 0$ such that both m and M oscillate at angular frequency ω in opposite directions while the center of mass does not shift.

12. Atwood's Machine with Massive String**

As the string is inextensible, the masses and all segments of the string must have the same speed. The Lagrangian of the system is

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2 + m_3)\dot{x}^2 + m_1 gx + m_2 g(l - x) + \frac{l^2 - 2lx + 2x^2}{2l}m_3 g.$$

The E-L equations yield

$$(m_1 + m_2 + m_3)\ddot{x} = m_1 g - m_2 g + \left(\frac{2x}{l} - 1\right)m_3 g.$$

Simplifying,

$$\ddot{x} = \frac{2gm_3}{(m_1 + m_2 + m_3)l}x + \frac{m_1 - m_2 - m_3}{m_1 + m_2 + m_3}g.$$

Using the substitution $x' = x + \frac{(m_1 - m_2 - m_3)l}{2m_3}$,

$$\ddot{x}' = \frac{2gm_3}{(m_1 + m_2 + m_3)l} x'.$$

The general solution to this differential equation is

$$\begin{aligned} x' &= Ae^{\sqrt{\frac{2gm_3}{(m_1+m_2+m_3)l}}t} + Be^{-\sqrt{\frac{2gm_3}{(m_1+m_2+m_3)l}}t} \\ \implies x &= Ae^{\sqrt{\frac{2gm_3}{(m_1+m_2+m_3)l}}t} + Be^{-\sqrt{\frac{2gm_3}{(m_1+m_2+m_3)l}}t} - \frac{(m_1 - m_2 - m_3)l}{2m_3}. \end{aligned}$$

Then, $\alpha = \sqrt{\frac{2gm_3}{(m_1+m_2+m_3)l}}$.

13. Brachistochrone**

The velocity of the bead at a y-coordinate y is given by the classical conservation of energy as

$$v = \sqrt{2gy}.$$

An infinitesimal length along the wire $y(x)$ is

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'^2}dx.$$

Therefore, the time that we are trying to minimize is

$$t = \int_A^B \frac{ds}{v} = \int_A^B \sqrt{\frac{1 + y'^2}{2gy}} dx,$$

with predetermined endpoints A and B. We can thus apply the results from the calculus of variations, with the "Lagrangian" being

$$\mathcal{L} = \sqrt{\frac{1 + y'^2}{2gy}}.$$

However, instead of applying the E-L equation, we can observe that the Lagrangian is independent of x. Therefore, the "Hamiltonian" is conserved.

Consequently, the function $y(x)$ that minimizes t satisfies

$$\mathcal{L} - \frac{\partial \mathcal{L}}{\partial y'} \cdot y' = c$$

for some constant c.

$$\sqrt{\frac{1+y'^2}{2gy}} - \frac{y'^2}{\sqrt{2gy(1+y'^2)}} = c$$

$$\frac{1}{2gy(1+y'^2)} = c^2$$

$$y' = \sqrt{\frac{\frac{1}{2gc^2} - y}{y}}.$$

Expressing $\frac{1}{2gc^2}$ as a new constant β,

$$y' = \sqrt{\frac{\beta - y}{y}}.$$

This suggests that we should try the parameterization $y = \beta \sin^2 \theta$ for some variable θ. Then,

$$y' = \cot \theta.$$

Since $y' = \frac{dy}{dx} = \frac{dy}{d\theta} \cdot \frac{d\theta}{dx}$ and $\frac{dy}{d\theta} = 2\beta \sin\theta \cos\theta$,

$$\frac{dx}{d\theta} = \frac{\frac{dy}{d\theta}}{y'} = 2\beta \sin^2 \theta.$$

Applying the formula $2\sin^2\theta = 1 - \cos 2\theta$,

$$\int_0^x dx = \int_0^\theta \beta(1 - \cos 2\theta) d\theta$$

$$x = \beta\left(\theta - \frac{\sin 2\theta}{2}\right).$$

Defining new variables $\phi = 2\theta$ and $a = \frac{\beta}{2}$,

$$x = a(\phi - \sin\phi),$$

$$y = a(1 - \cos\phi).$$

The curve described by these parametric equations is known as a cycloid and its geometric interpretation is given in the solution to the next problem.

14. Wire Pendulum**

The kinetic energy of the particle is $\frac{1}{2}m\dot{s}^2$ while its gravitational potential energy is $mgy = mgks^2$. Therefore, the Lagrangian of the particle is

$$\mathcal{L} = \frac{1}{2}m\dot{s}^2 - mgks^2.$$

Applying the E-L equation,

$$m\ddot{s} = -2mgks$$
$$\ddot{s} = -2gks,$$

which indicates a simple harmonic motion. This is independent of the initial coordinate and velocity of the particle along the wire as we have not made any assumptions about them (e.g. that s and $\dot{s}$ are small). The definition of $s(x)$ implies

$$s(x) = \int_0^x \sqrt{1+y'^2}dx.$$

Therefore, $\frac{ds}{dx} = \sqrt{1+y'^2}$. Differentiating $y = ks^2$ with respect to x,

$$y' = 2ks\sqrt{1+y'^2}$$
$$\frac{y'}{\sqrt{1+y'^2}} = 2ks = 2\sqrt{ky}.$$

The above equation suggests the substitution $y' = \tan\theta$ such that

$$\sin\theta = 2\sqrt{ky}$$
$$y = \frac{\sin^2\theta}{4k}.$$

We can now solve for $x(\theta)$ by using the fact that

$$y' = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = \frac{\frac{\sin\theta\cos\theta}{2k}}{\frac{dx}{d\theta}},$$
$$\frac{dx}{d\theta} = \frac{\cos^2\theta}{2k} = \frac{\cos 2\theta + 1}{4k}$$
$$\implies x = \frac{\sin 2\theta}{8k} + \frac{\theta}{4k} + c,$$

where c is a constant. Notice that the origin must correspond to $\theta = 0$ as $y' = \tan\theta = 0$ at $x = 0$ (as required by the wire being horizontal at the origin). Therefore, $x(0)$ must be zero — implying that

$$c = 0,$$
$$x = \frac{\sin 2\theta}{8k} + \frac{\theta}{4k},$$
$$y = \frac{\sin^2 \theta}{4k} = \frac{1}{8k} - \frac{1}{8k}\cos 2\theta.$$

We can rewrite the above in a more suggestive form by introducing a new variable $\phi = 2\theta$ such that

$$x = \frac{1}{8k}\sin\phi + \frac{\phi}{8k},$$
$$y = \frac{1}{8k} - \frac{1}{8k}\cos\phi.$$

Observe that the above is akin to the trajectory of a particle attached to the rim of a circle of radius $\frac{1}{8k}$, rolling without slipping in the positive x-direction along the ground at $y = 0$. The particle begins at the bottom of the circle. The center of the circle after the circle has rotated anti-clockwise by angle ϕ is at $(\frac{\phi}{8k}, \frac{1}{8k})$ — implying that the coordinates of the particle are described by the above equations. Such a trajectory is known as a cycloid. Finally, notice that the curve described in the solution to the previous problem is also a cycloid. It is remarkable how ubiquitous a cycloid is (and how it often optimizes something). In fact, we will encounter a cycloid again when we analyze the motion of a charge in mutually perpendicular and constant electric and magnetic fields.

15. Virial Theorem**

Firstly, G must be bounded (i.e. does not diverge) as the coordinates and velocities of the particles are bounded. Therefore, $\int_0^t \frac{dG}{dt}dt = G(t) - G(0)$ must be finite for all times t — implying that

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \frac{dG}{dt}dt = 0.$$

Now, let us evaluate $\frac{dG}{dt}$ explicitly.

$$\frac{dG}{dt} = \sum_{i=1}^{N} m_i \dot{\boldsymbol{r}}_i \cdot \dot{\boldsymbol{r}}_i + \sum_{i=1}^{N} m_i \boldsymbol{r}_i \cdot \ddot{\boldsymbol{r}}_i.$$

The first summation is simply equal to twice the total kinetic energy of the system, $2T$. Meanwhile, since $m_i\ddot{\boldsymbol{r}}_i$ is equal to the force experienced by the

ith particle, $-\frac{\partial U}{\partial \boldsymbol{r}_i}$, the above becomes

$$\begin{aligned}\frac{dG}{dt} &= 2T - \sum_{i=1}^{N} \boldsymbol{r}_i \cdot \frac{\partial U}{\partial \boldsymbol{r}_i} \\ &= 2T - \sum_{i=1}^{N} \left(x_i \frac{\partial U}{\partial x_i} + y_i \frac{\partial U}{\partial y_i} + z_i \frac{\partial U}{\partial z_i} \right) \\ &= 2T - dU,\end{aligned}$$

where we have applied Euler's theorem of homogeneous functions in the last step. Taking the time-average of both sides over an infinite period,

$$\langle 2T - dU \rangle = 0.$$

In fact, the virial theorem is extremely useful in estimating the total mass in a galaxy (which can be safely assumed to be bounded since it has been there for a long time). For a system under a gravitational potential, $d = -1$ such that $\langle 2T + U \rangle = 0$. The total kinetic energy can be estimated via the Doppler effect while the gravitational potential energy can be approximated in a manner similar to that of a spherical blob (with a measurable radius) such that the total mass in the galaxy can be estimated. In fact, this method led to the earliest hypothesis about the existence of dark matter, as the mass of the Milky Way galaxy, predicted by this model, vastly differed from the total mass of the observable stars.

16. Cylinder on Cylinder***

Let the angles that the bottom and top cylinders have rotated about their own centers be ψ and ϕ respectively. They are measured clockwise from the vertical axis. Define the origin to be at the center of the bottom cylinder. Let the radial coordinate of the center of the top cylinder be r (we set this as a variable as we wish to include a Lagrange multiplier). Then, $r\dot{\theta}$ is the azimuthal velocity of the center of the top cylinder. Hence, the non-slip condition is

$$r\dot{\theta} - r_2\dot{\phi} = r_1\dot{\psi}.$$

Furthermore, the condition for the top cylinder to remain on the surface of the bottom cylinder is

$$r = r_1 + r_2.$$

Now, the Lagrangian of the system comprising the two cylinders will be written down. The moments of inertia are $\frac{1}{2}Mr_1^2$ and $\frac{1}{2}mr_2^2$ respectively.

Hence, the Lagrangian of the system with the Lagrange multiplier is

$$\mathcal{L} = \frac{1}{4}Mr_1^2\dot{\psi}^2 + \frac{1}{4}mr_2^2\dot{\phi}^2 + \frac{1}{2}m(r^2\dot{\theta}^2 + \dot{r}^2) - mgr\cos\theta + \lambda(r - r_1 - r_2).$$

Substituting the expression for $\dot{\psi}$ obtained from the non-slip condition,

$$\mathcal{L} = \frac{1}{4}M(r\dot{\theta} - r_2\dot{\phi})^2 + \frac{1}{4}mr_2^2\dot{\phi}^2 + \frac{1}{2}m(r^2\dot{\theta}^2 + \dot{r}^2) - mgr\cos\theta + \lambda(r - r_1 - r_2).$$

Applying the E-L equation with respect to r,

$$m\ddot{r} = \frac{1}{2}M\left(r\dot{\theta} - r_2\dot{\phi}\right)\dot{\theta} + mr\dot{\theta}^2 - mg\cos\theta + \lambda. \tag{12.12}$$

For the rest of the coordinates θ and ϕ, r can be treated as a constant when applying the E-L equation as the Lagrange multiplier term does not depend on θ and ϕ.

$$\frac{1}{2}M(r\ddot{\theta} - r_2\ddot{\phi})r + mr^2\ddot{\theta} = mgr\sin\theta, \tag{12.13}$$

$$\frac{1}{2}M(r\ddot{\theta} - r_2\ddot{\phi})r_2 + \frac{1}{2}mr_2^2\ddot{\phi} = 0. \tag{12.14}$$

Recall that our ultimate objective is to solve for $\lambda\frac{\partial(r-r_1-r_2)}{\partial r} = \lambda$, which represents the normal force, in terms of θ. Hence, we have to eliminate all other variables. From Eq. (12.14),

$$\ddot{\phi} = \frac{Mr}{(M-m)r_2}\ddot{\theta}.$$

Substituting this expression for $\ddot{\phi}$ in Eq. (12.13),

$$\ddot{\theta} = \frac{2(M-m)g}{(M-2m)r}\sin\theta.$$

This is the expression for $\ddot{\theta}$ in terms of θ. Now, we wish to solve for $\dot{\theta}$ which appears in Eq. (12.12). By using the fact that $\ddot{\theta} = \frac{d\dot{\theta}}{d\theta}\dot{\theta}$, shifting $d\theta$ to the right-hand side and integrating both sides, we obtain

$$\int_0^{\dot{\theta}} \dot{\theta}d\dot{\theta} = \int_0^{\theta} \frac{2(M-m)g}{(M-2m)r}\sin\theta d\theta$$

$$\dot{\theta}^2 = \frac{4(M-m)g}{(M-2m)r}(1-\cos\theta).$$

By substituting $\dot{\phi} = \frac{Mr}{(M-m)r_2}\dot{\theta}$ in Eq. (12.12) and enforcing the constraint $\ddot{r} = 0$,

$$\lambda = mg\cos\theta - \frac{m(M-2m)r}{2(M-m)}\dot{\theta}^2.$$

Substituting the expression for $\dot{\theta}^2$,

$$\lambda = 3mg\cos\theta - 2mg.$$

The normal force is

$$\lambda\frac{\partial(r - r_1 - r_2)}{\partial r} = \lambda.$$

Hence, the top cylinder loses contact with the bottom when $\lambda = 0$. Then,

$$\theta_{max} = \cos^{-1}\left(\frac{2}{3}\right).$$

17. Inverted Pendulum***

The coordinates of the particle are

$$(x, y) = (l\sin\theta, A\cos\omega t + l\cos\theta) \implies (\dot{x}, \dot{y}) = (l\cos\theta\dot{\theta}, -A\omega\sin\omega t - l\sin\theta\dot{\theta}).$$

The Lagrangian of the particle is thus (after discarding the total time derivatives)

$$\mathcal{L} = \frac{1}{2}m(l^2\dot{\theta}^2 + 2Al\omega\sin\omega t\sin\theta\dot{\theta}) - mgl\cos\theta.$$

Applying the E-L equation,

$$ml^2\ddot{\theta} + mAl\omega^2\cos\omega t\sin\theta + mAl\omega\sin\omega t\cos\theta\dot{\theta} = mgl\sin\theta.$$

For small θ and $\dot{\theta}$,

$$\ddot{\theta} + (a\omega^2\cos\omega t - \omega_0^2)\theta = 0,$$

where $a = \frac{A}{l} \ll 1$ and $\omega_0 = \sqrt{\frac{g}{l}}$. When $a\omega^2 \gg \omega_0^2$ (this shall be our definition of large ω for now), θ effectively oscillates with the platform at angular frequency ω, virtually neglecting gravity, during time-scales considerably shorter than $\frac{2\pi}{\omega_0}$ such that we can try a solution

$$\theta = B + C\cos\omega t,$$

where the time-scales of B and C are much larger than $\frac{2\pi}{\omega}$. Furthermore, $B \gg C$ as $A \ll l$ such that the effect of the platform on the position of m

is small in a single period. Substituting the above expression for θ into the previous equation while keeping B and C constant,

$$-C\omega^2 \cos\omega t + a\omega^2 \cos\omega t B = 0$$
$$\implies C = aB,$$
$$\theta = B(1 + a\cos\omega t).$$

Substituting this expression into $\ddot{\theta} + (a\omega^2 \cos\omega t - \omega_0^2)\theta = 0$,

$$\langle\ddot{\theta}\rangle = -\langle(a\omega^2 \cos\omega t - \omega_0^2)(1 + a\cos\omega t)B\rangle = -\left(\frac{a^2\omega^2}{2} - \omega_0^2\right)B,$$

where we average over the relatively short period $\frac{2\pi}{\omega}$. Next, we also know from $\theta = B(1 + a\cos\omega t)$ that

$$\begin{aligned}
\dot{\theta} &= \dot{B}(1 + a\cos\omega t) - a\omega B \sin\omega t \\
\ddot{\theta} &= \ddot{B}(1 + a\cos\omega t) - 2a\omega\dot{B}\sin\omega t - a\omega^2 B\cos\omega t \\
\implies \langle\ddot{\theta}\rangle &= \ddot{B}, \\
\implies \ddot{B} &= -\left(\frac{a^2\omega^2}{2} - \omega_0^2\right)B,
\end{aligned}$$

which indicates a simple harmonic motion of angular frequency

$$\Omega = \sqrt{\frac{a^2\omega^2}{2} - \omega_0^2} = \sqrt{\frac{A^2\omega^2}{2l^2} - \frac{g}{l}}$$

if $a\omega > \sqrt{2}\omega_0$ (implying that our initial assumption $a\omega^2 \gg \omega_0^2$ needs to be further strengthened since $a \ll 1$). Therefore,

$$\theta \approx D\sin(\Omega t + \phi)(1 + a\cos\omega t)$$

for some constants D and ϕ. This is oscillatory[4] and is bounded by a maximum value $D(1 + a)$ which must be small when its initial angular deviation from the vertical is small. Therefore, the stick does not topple.

18. Double Pendulum***

Define the origin to be at the fixed end of the first string. Then the coordinates of m_1 and m_2 are $(l_1 \sin\theta, -l_1\cos\theta)$ and $(l_1\sin\theta_1 +$

[4]As $\Omega \approx \frac{a}{\sqrt{2}}\omega \ll \omega$, we can treat $\sin(\Omega t + \phi)$ as a constant over the period of $\cos\omega t$.

$l_2 \sin\theta_2, -l_1\cos\theta_1 - l_2\cos\theta_2)$ respectively. The squared speed of m_2 can be obtained from differentiating its coordinates to be

$$\begin{aligned}&\left(l_1\cos\theta_1\dot{\theta}_1 + l_2\cos\theta_2\dot{\theta}_2\right)^2 + \left(l_1\sin\theta_1\dot{\theta}_1 + l_2\sin\theta_2\dot{\theta}_2\right)^2\\ &\quad = l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2).\end{aligned}$$

The Lagrangian of the two bobs is then

$$\begin{aligned}\mathcal{L} = &\frac{1}{2}m_1l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2(l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1l_2\dot{\theta}_1\dot{\theta}_2\cos(\theta_1-\theta_2))\\ &+ m_1gl_1\cos\theta_1 + m_2g(l_1\cos\theta_1 + l_2\cos\theta_2).\end{aligned}$$

The E-L equations with respect to θ_1 and θ_2 yield

$$\begin{aligned}&(m_1+m_2)l_1^2\ddot{\theta}_1 + m_2l_1l_2\ddot{\theta}_2\cos(\theta_1-\theta_2) - m_2l_1l_2\dot{\theta}_2\sin(\theta_1-\theta_2)(\dot{\theta}_1-\dot{\theta}_2)\\ &\quad = -m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1-\theta_2) - (m_1+m_2)gl_1\sin\theta_1,\\ &m_2l_2^2\ddot{\theta}_2 + m_2l_1l_2\ddot{\theta}_1\cos(\theta_1-\theta_2) - m_2l_1l_2\dot{\theta}_1\sin(\theta_1-\theta_2)(\dot{\theta}_1-\dot{\theta}_2)\\ &\quad = m_2l_1l_2\dot{\theta}_1\dot{\theta}_2\sin(\theta_1-\theta_2) - m_2gl_2\sin\theta_2.\end{aligned}$$

Applying small angle approximations and discarding second-order terms ($\dot{\theta}_1$ and $\dot{\theta}_2$ are considered to be small too), we obtain the following after some simplification.

$$\begin{aligned}(m_1+m_2)l_1\ddot{\theta}_1 + (m_1+m_2)g\theta_1 + m_2l_2\ddot{\theta}_2 &= 0,\\ l_1\ddot{\theta}_1 + l_2\ddot{\theta}_2 + g\theta_2 &= 0.\end{aligned}$$

This set of equations represents the motion of coupled oscillators. Using the substitution $\theta_1 = Ae^{i\omega t}$ and $\theta_2 = Be^{i\omega t}$ and expressing the resultant set of equations in matrix form,

$$\begin{pmatrix}(m_1+m_2)(l_1\omega^2 - g) & m_2l_2\omega^2\\ l_1\omega^2 & l_2\omega^2 - g\end{pmatrix}\begin{pmatrix}A\\ B\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}.$$

For non-trivial solutions to exist for A and B, the matrix on the left must be singular — implying that its determinant must be zero. Then,

$$(m_1+m_2)(l_1\omega^2 - g)(l_2\omega^2 - g) - m_2l_2\omega^2\cdot l_1\omega^2 = 0.$$

Solving for the possible values of ω yields

$$\omega_\pm = \sqrt{\frac{(m_1+m_2)(l_1+l_2)g \pm g\sqrt{(l_1-l_2)^2m_1^2 + (l_1+l_2)^2m_2^2 + 2(l_1^2+l_2^2)m_1m_2}}{2m_1l_1l_2}}.$$

19. Unraveling Rope***

Define the origin to be at the support and let the x-coordinate of the center of mass of the spiral be x. Then, the mass m and radius r of the remaining spiral as functions of x are

$$m = M\left(1 - \frac{x}{L}\right),$$
$$r = R\left(1 - \frac{x}{L}\right).$$

Since the spiral is akin to a circular disk, its moment of inertia is

$$I = \frac{1}{2}mr^2.$$

Following from this, its rotational kinetic energy, given that its instantaneous angular velocity is ω, is

$$\frac{1}{2}I\omega^2 = \frac{1}{4}mr^2\omega^2 = \frac{1}{4}m\dot{x}^2,$$

where we have applied the non-slip condition $r\omega = \dot{x}$. Therefore, the Lagrangian of the entire rope, whose sole contributor is the remaining portion of the spiral, is

$$\begin{aligned}\mathcal{L} &= \frac{1}{2}m\dot{x}^2 + \frac{1}{2}I\omega^2 - mgr \\ &= \frac{3}{4}M\left(1 - \frac{x}{L}\right)\dot{x}^2 - MgR\left(1 - \frac{x}{L}\right)^2.\end{aligned}$$

Computing the relevant derivatives,

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial x} &= -\frac{3M}{4L}\dot{x}^2 + \frac{2MgR}{L}\left(1 - \frac{x}{L}\right), \\ \frac{\partial \mathcal{L}}{\partial \dot{x}} &= \frac{3}{2}M\left(1 - \frac{x}{L}\right)\dot{x}, \\ \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) &= -\frac{3M}{2L}\dot{x}^2 + \frac{3}{2}M\left(1 - \frac{x}{L}\right)\ddot{x}.\end{aligned}$$

Applying the E-L equation and equating the first and third expressions while using $\ddot{x} = \frac{d\dot{x}^2}{2dx}$,

$$\frac{3}{4}M\left(1 - \frac{x}{L}\right)\frac{d\dot{x}^2}{dx} - \frac{3M}{4L}\dot{x}^2 = \frac{2MgR}{L}\left(1 - \frac{x}{L}\right).$$

Simplifying,

$$(L-x)\frac{d\dot{x}^2}{dx}-\dot{x}^2=\frac{8gR}{3L}(L-x).$$

Observe that the left-hand side is the total derivative $\frac{d[(L-x)\dot{x}^2]}{dx}$. Separating variables and integrating,

$$\int d\left[(L-x)\dot{x}^2\right]=\int_0^x \frac{8gR}{3L}(L-x)dx.$$

Setting the constant of integration of the left-hand side as $-\frac{4gR}{3L}c$ where c is a positive constant (note that the constant of integration must be negative as it is negative of L times the initial $\dot{x}$ squared),

$$(L-x)\dot{x}^2-\frac{4gR}{3L}c=\frac{4gR}{3L}(2xL-x^2).$$

Rearranging,

$$\dot{x}^2=\frac{4gR}{3L(L-x)}(2xL-x^2+c).$$

The only way for the rope to stop unraveling is for $\dot{x}$ to attain zero at some point before $x=L$ so that it is possible for $\dot{x}$ to become negative or remain at zero. This requires

$$x^2-2xL-c=0,$$

with solutions

$$x=\frac{2L\pm\sqrt{4L^2+4c}}{2}=L\pm\sqrt{L^2+c}.$$

Notice that both solutions do not satisfy the above requirement as one is negative while the other is larger than L. Therefore, the rope must completely unravel. Moving on, in the case where the initial $\dot{x}$ is negligible, $c=0$. Then,

$$\dot{x}=\sqrt{\frac{4gR}{3L}\cdot\frac{x(2L-x)}{L-x}}.$$

To determine the time t required for the rope to completely unravel, we separate variables and integrate the above from $x=0$ to $x=L$:

$$\int_0^L\sqrt{\frac{L-x}{x(2L-x)}}dx=\int_0^t\sqrt{\frac{4gR}{3L}}dt.$$

To simplify the integral on the left, we use the trick

$$\int_a^b f(x)dx = \int_a^b f(a+b-x)dx,$$

which applies to any definite integral. In this case, $a = 0$ and $b = L$ so we substitute $L - x$ for x in the integrand and get

$$\int_0^L \sqrt{\frac{L-x}{x(2L-x)}}dx = \int_0^L \sqrt{\frac{x}{L^2-x^2}}dx.$$

Now, we can adopt the trigonometric substitution $x = L\sin\theta$ and $dx = L\cos\theta d\theta$.

$$\int_0^L \sqrt{\frac{x}{L^2-x^2}}dx = \int_0^{\frac{\pi}{2}} \sqrt{L\sin\theta}d\theta = \sqrt{L}\alpha,$$

where we have agreed on using α to denote $\int_0^{\frac{\pi}{2}} \sqrt{\sin\theta}d\theta$. Returning to the original equation, we have

$$\sqrt{L}\alpha = \sqrt{\frac{4gR}{3L}}t.$$

Therefore,

$$t = \frac{\sqrt{3}L\alpha}{\sqrt{4gR}}.$$

This page intentionally left blank

Chapter 13

Waves

In this chapter, we will analyze waves which form an important concept in various fields of physics. No new laws are introduced in this chapter. Mechanical waves are still studied via the equation $F = ma$ and electromagnetic waves can be investigated by considering Maxwell's equations. Waves can be thought to be the mathematical consequence of the physical laws, though they are surprisingly common and have profound physical meaning.

In this chapter, we will only be analyzing non-dispersive waves, whose speed of propagation is independent of the wavelength of the wave.

13.1 Introduction

Consider a rope that is held under tension. If we give the left-end of the rope a little wiggle, a "pulse" (known as a waveform) will be sent across the rope.

The string segment on the immediate right of the initial waveform will be pulled upwards by the tension of the pulse. Then the next adjacent string segment is also pulled upwards by the new shifted waveform and the wave propagates through the string, resulting in a transfer of energy. However, each section of the rope only oscillates up and down about its equilibrium position and no section is actually physically transported from one end to another. This can be best visualized by marking a section of the string with a dot such as point P shown in Fig. 13.1.

This is an example of a traveling wave, which is a disturbance that is able to transmit energy or momentum from a source to its surroundings. In the case of a mechanical wave, the points in the medium that carry the wave do not actually travel through the region through which the wave propagates.

A wave is analogous to an infinite series of coupled oscillators, though the exact interactions between adjacent oscillators depend on the nature of the

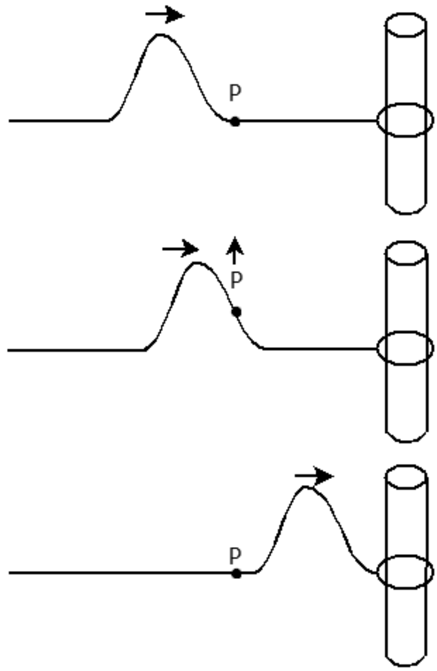


Figure 13.1: Traveling pulse

wave. Each point on the wave oscillates about its own equilibrium position. In the case of a finite number of oscillators, each oscillator is associated with its own displacement which is a function of time. As a wave can be seen as an array of an infinite number of oscillators, the displacement $\psi(x,t)$ is used to denote the displacement of a point at the x-coordinate x on a one-dimensional wave from its equilibrium position at time t.

13.1.1 *Nature of Waves*

There are two main types of waves that are of interest to us — namely mechanical and electromagnetic waves. Definitely, these do not form an exhaustive list. There are also other types of waves such as matter waves and gravitational waves.

Mechanical Waves

A mechanical wave involves oscillations of matter. Mechanical waves propagate due to restoring forces on particles that are displaced from their equilibrium positions. For instance, the disproportionate pressure on two sides of a section of air molecules provides the restoring force in a sound wave. Therefore, all mechanical waves require media with mass and elasticity, whose properties determine their speed of propagation, for propagation. Common examples of mechanical waves include string waves, sound waves and water waves.

Electromagnetic Waves (EM Waves)

An electromagnetic wave consists of an electric and magnetic field, both time-varying, which are perpendicular both to the direction of wave propagation, as well as to each other. As opposed to the movement of sections of string as a string wave propagates, an electromagnetic wave leads to changes of the surrounding electric and magnetic fields as the EM wave propagates.

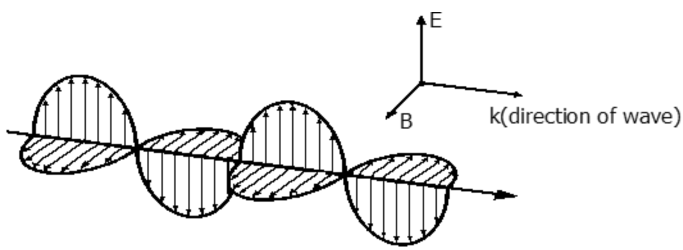


Figure 13.2: Electric and magnetic field displacements in an EM wave

An electromagnetic wave propagates due to a self-sustaining induction mechanism. A changing magnetic field induces a changing electric field and vice-versa — the variations of the two fields are interconnected. As the propagation of an electromagnetic wave relies solely on electromagnetic induction, an EM wave is able to travel in a vacuum. In fact, all EM waves travel at speed c in vacuum, which is commonly known as the speed of light in vacuum.

$$c = 3.00 \times 10^8 m/s.$$

13.1.2 *Direction of Vibration*

Waves can also be divided into longitudinal and transverse waves based on the direction of oscillation of points on the waves. A transverse wave is one in which points of disturbance oscillate about their equilibrium positions perpendicular to the direction of wave propagation. All EM waves are transverse waves. Transverse waves also have crests and troughs which correspond to the points of the medium that have the most positive and negative displacements from their equilibrium positions at a certain instant in time.

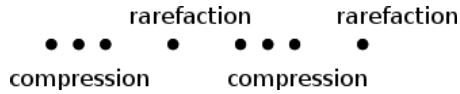


Figure 13.3: Regions of compression and rarefaction

A longitudinal wave is one in which points of disturbance oscillate about their equilibrium positions in a direction parallel to that of wave propagation. A sound wave is an example of a longitudinal wave. Rather than crests and troughs, a longitudinal wave contains regions where adjacent particles are the closest and furthest with respect to each other at a certain instant. These regions are known as points of compression and rarefaction, respectively.

In a sound wave, points of compression and rarefaction (Fig. 13.3) at a particular instant in time correspond to the points of maximum and minimum pressure at that instant, respectively. Furthermore, they both correspond to points of zero displacement as a particle at such points must be at its equilibrium position in order to be squeezed by or evacuated of surrounding molecules to the greatest extent.

Some waves, such as surface water waves, are a combination of both transverse and longitudinal waves and are assigned to neither of the above categories.

13.1.3 *Definitions*

Certain terminologies are used to describe a wave. The diagram below depicts a "snapshot" of a one-dimensional traveling sinusoidal wave at a particular instant. The y-axis corresponds to the displacements of particles from their equilibrium position at a particular instant in time, which may be in the transverse or longitudinal direction.

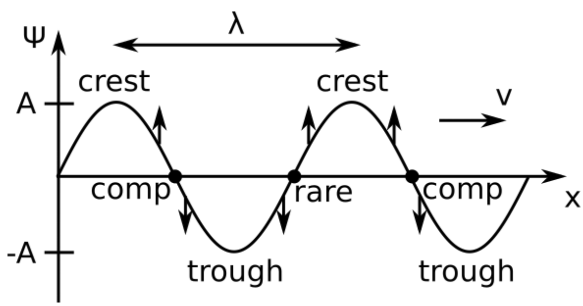


Figure 13.4: Traveling sinusoidal wave

- The displacement, $\psi(x, t)$, of a wave refers to the distance and direction of the particle point at x-coordinate, x, from its equilibrium position at time t. Note that ψ depends on both x and t. Figure 13.4 only shows the displacements of the points on the wave at a particular instant. At the next instant, the entire wave profile will shift to the right and the displacements of the points will change.
- The amplitude, A, of a periodic wave at coordinate x is the maximum magnitude of the displacement of the point on the wave at coordinate x over all possible times. In the case of the sinusoidal wave depicted, the amplitudes at all points on the wave are the same.
- The wavelength λ of a periodic wave is the distance between two adjacent points that are in the same state of oscillation (with a phase difference of 2π) at a particular instant. This can be computed as the distance between two successive crests or troughs for transverse waves and between two successive compressions or rarefactions in the case of a longitudinal wave. The labels in the figure are technically incorrect as the displacement cannot refer to both the transverse and longitudinal displacements at the same time. They are superimposed for the sake of simplicity.

  One important fact to understand is that the displacement is along the direction of propagation for longitudinal waves (i.e. the particles in the above wave are displaced along the x-direction) though it is seemingly transverse. To discern between compressions and rarefactions which are both reflected by equilibrium points in a displacement graph, one has to consider the instantaneous velocities of surrounding points. For example, the first equilibrium point labeled corresponds to a point of compression as its left neighbor has a positive velocity (towards larger x as the displacement is longitudinal) while its left neighbor has a negative velocity. To visualize these velocities, simply consider the displacements of those particles when the wave shifts slightly rightwards (in the direction of propagation). Evidently, the particle at this equilibrium position is squeezed from both sides and hence "compressed." Similarly, the neighboring molecules of the second equilibrium point labeled diverge from the equilibrium point — implying that it corresponds to a point of rarefaction.
- The period, T, of a periodic wave is the time taken for a point on the wave to complete one oscillation cycle. That is, we fix a point of consideration and observe how long it takes to return to its initial state. The period is also the time taken for the wave to advance by a distance of one wavelength as that is how long it takes for the corresponding point of the wave to travel from another location to the point under consideration.

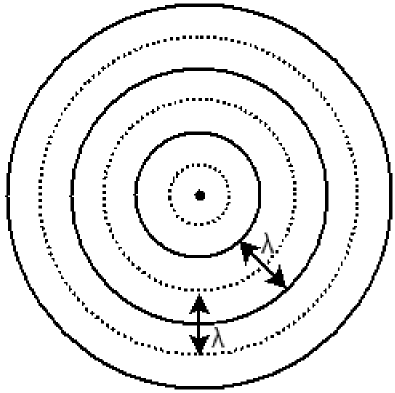


Figure 13.5: Ripples

- The frequency, f, of a periodic wave is then the number of oscillation cycles completed by a point on the wave per unit time, $f = \frac{1}{T}$.
- The phase angle of a traveling sinusoidal wave at a certain point in space represents the state of oscillation of that point at a particular time. For example, for the sinusoidal traveling wave with $\psi(x, t)$ given by

$$\psi = A \cos(kx - \omega t + \phi),$$

  the quantity $kx - \omega t + \phi$ is the phase angle of the point on the wave at coordinate x at time t.
- A wavefront of a periodic wave is a locus of points on the wave that are at the same state of oscillation (in phase) at a particular point in time. By convention, the distance between consecutive wave fronts is drawn to correspond to one wavelength. A vivid and intuitive illustration of wave fronts would be the ripples that are obtained when a stone is dropped into a pond (Fig. 13.5).
  The crests of the water waves are joined to form wave fronts that are delineated by solid lines while the troughs are connected to form wave fronts that are represented by dotted lines. A trough lies in the middle of two consecutive crests and vice-versa.
- The phase velocity, v, of a wave is the distance that the wave profile appears to traverse per unit time. Since a periodic wave travels a distance λ in a time period T, the phase velocity of a traveling wave satisfies the relationship

$$v = \frac{\lambda}{T} = f\lambda. \tag{13.1}$$

Note that except for displacement and the speed of a wave, all quantities in the bullet points above are usually defined only for waves with continuous and periodic energy input. For example, a person would have to keep wiggling a string under tension or continue dropping stones in a pond to maintain string and water waves, respectively. If the waves are not sustained, there will only be a limited number of "pulses" that are transmitted. Then, the analyses of such waves are much harder to quantify.

13.2 The Wave Equation

To delve further into the concept of a wave, one has to turn to the quantitative formulation of a wave. All one-dimensional waves along the x-direction are described by the following partial differential equation, known as the one-dimensional wave equation.

$$\frac{\partial^2\psi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}, \tag{13.2}$$

where $\psi(x,t)$ refers to the displacement of a certain quantity from its equilibrium value at x-coordinate x and time t. It could be the displacement of a certain particle on a string at a certain point in time or the electric field at a certain spatial position at a particular time. v is a constant that we will discover to be the phase velocity of the wave.

All types of waves boil down to equations of this form, though the exact mechanisms involved in their derivations may differ across various waves. Any quantity ψ that satisfies the above partial differential equation corresponds to a wave.

13.2.1 *General Solution*

The general solution to the one-dimensional wave equation can be obtained elegantly by D'Alembert's method. We define new variables $\xi = x - vt$ and $\eta = x + vt$. We can rewrite $\psi(x,t)$ as a function of ξ and η easily, i.e.

$$\psi(x,t) = \psi\left(\frac{\xi+\eta}{2}, \frac{\eta-\xi}{2v}\right) = \psi(\xi,\eta).$$

Then the partial derivatives in the wave equation can be rewritten as

$$\begin{aligned}\frac{\partial}{\partial x} &= \frac{\partial}{\partial \xi}\cdot\frac{\partial \xi}{\partial x} + \frac{\partial}{\partial \eta}\cdot\frac{\partial \eta}{\partial x}\\ &= \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta},\end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial t} &= \frac{\partial}{\partial \xi} \cdot \frac{\partial \xi}{\partial t} + \frac{\partial}{\partial \eta} \cdot \frac{\partial \eta}{\partial t} \\
&= -v \frac{\partial}{\partial \xi} + v \frac{\partial}{\partial \eta}, \\
\frac{\partial^2}{\partial x^2} &= \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \right) \left(\frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \right) \\
&= \frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \xi \partial \eta} + \frac{\partial^2}{\partial \eta \partial \xi} + \frac{\partial^2}{\partial \eta^2} \\
&= \frac{\partial^2}{\partial \xi^2} + 2 \frac{\partial^2}{\partial \xi \partial \eta} + \frac{\partial^2}{\partial \eta^2}, \\
\frac{\partial^2}{\partial t^2} &= \left(-v \frac{\partial}{\partial \xi} + v \frac{\partial}{\partial \eta} \right) \left(-v \frac{\partial}{\partial \xi} + v \frac{\partial}{\partial \eta} \right) \\
&= v^2 \frac{\partial^2}{\partial \xi^2} - v^2 \frac{\partial^2}{\partial \xi \partial \eta} - v^2 \frac{\partial^2}{\partial \eta \partial \xi} + v^2 \frac{\partial^2}{\partial \eta^2} \\
&= v^2 \frac{\partial^2}{\partial \xi^2} - 2v^2 \frac{\partial^2}{\partial \xi \partial \eta} + v^2 \frac{\partial^2}{\partial \eta^2}.
\end{aligned}
$$

The third equalities of the last two expressions stem from the fact that the order in which partial derivatives are taken does not matter. Substituting these partial derivatives into the wave equation, we obtain

$$
\begin{aligned}
\frac{\partial^2 \psi}{\partial \xi^2} + 2 \frac{\partial^2 \psi}{\partial \xi \partial \eta} + \frac{\partial^2 \psi}{\partial \eta^2} &= \frac{1}{v^2} \left(v^2 \frac{\partial^2 \psi}{\partial \xi^2} - 2v^2 \frac{\partial^2 \psi}{\partial \xi \partial \eta} + v^2 \frac{\partial^2 \psi}{\partial \eta^2} \right) \\
\implies \frac{\partial^2 \psi}{\partial \xi \partial \eta} &= 0.
\end{aligned}
$$

We can integrate the expression above with respect to η first. The constant of integration is a function of ξ that is independent of η, as we treat it as a constant when taking the partial derivative of $\frac{\partial \psi}{\partial \xi}$ with respect to η (remember that ψ is now a function of ξ and η). Thus, we obtain

$$
\frac{\partial \psi}{\partial \xi} = h(\xi)
$$

for some function $h(\xi)$. Integrating this expression with respect to ξ and letting the constant of integration be $g(\eta)$, an arbitrary function of η that is independent of ξ, we obtain

$$\begin{aligned}\psi &= \int h(\xi)d\xi + g(\eta) \\ &= f(\xi) + g(\eta) \\ &= f(x - vt) + g(x + vt),\end{aligned}$$

where $f(x-vt)$ and $g(x+vt)$ are arbitrary functions with $(x-vt)$ and $(x+vt)$ as their respective arguments. From this general solution, it is evident that v is the phase velocity of the wave. Consider an arbitrary function, $f(x-vt)$, at a certain time t_0. We know that if we reduce the argument by a certain Δx, the function simply shifts towards the positive x-direction by that amount. This is exactly what occurs after a time interval of Δt has passed as the function becomes $f(x - v(t_0 + \Delta t))$. It can be seen that $f(x - vt)$ represents a wave form that is traveling in the positive x-direction. In time Δt, the wave profile appears to travel a distance $v\Delta t$. Hence, v must refer to the phase velocity of the wave. A similar argument can be made for $g(x + vt)$ which represents a wave form traveling in the negative x-direction.

To obtain a more specific solution for ψ, it is necessary to impose initial conditions for ψ, such as the shape of the wave when $t = 0$, $\psi(x, 0)$, the initial velocity of different points of the wave, $\frac{\partial \psi}{\partial t}(x, 0)$, and possibly, other boundary conditions that we will explore later.

13.2.2 *One-Dimensional Traveling Sinusoidal Waves*

The particular solution for a sinusoidal progressive wave with a continuous energy input and no energy loss during propagation in the positive x-direction can be deduced from the above general solution to be

$$\psi = A \cos\left(k(x - vt) + \phi\right),$$

where ϕ is the phase offset of the wave and A is the amplitude of the wave. The boundary condition, in this case, is actually imposed by the fact that at all times, ψ must be a sinusoidal function of x that is traveling in the positive x-direction — the displacement at the driver of the wave (e.g. at the location of the hand wiggling the string) must itself be sinusoidal. Now, there is a need to multiply $(x - vt)$ by a quantity k, as the argument for a trigonometric function must be dimensionless. It is more edifying to express the above equation as

$$\psi = A \cos(kx - \omega t + \phi), \tag{13.3}$$

where ω is the angular frequency of oscillations of a point on the wave.

$$\omega = kv; \tag{13.4}$$

k is known as the wave number. By definition, the wavelength λ is the distance between two adjacent points that are oscillating in phase at a certain time. Thus, an additional λ increase in x should correspond to an increase in the argument in the cosine function of 2π. Therefore,

$$k = \frac{2\pi}{\lambda}. \tag{13.5}$$

Sometimes, it may be convenient to represent the displacement $\psi(x, t)$ of a one-dimensional traveling sinusoidal wave in terms of the real component of a complex displacement $\tilde{\psi}$, so that

$$\tilde{\psi} = Ae^{i(kx-\omega t+\phi)}$$

for a wave traveling in the positive x-direction such that

$$\psi = \mathrm{Re}(\tilde{\psi}).$$

The advantage of this formulation is that linear operations of $\tilde{\psi}$ can be performed in replacement of ψ — a feat that is often less tedious. The final physical wave can then be obtained from taking the real component of $\tilde{\psi}$.

Ultimately, the notion of sinusoidal waves is pivotal in the analysis of waves even though not all wave forms may take the form of a pure sinusoidal wave. This is because, any reasonably smooth function can be approximated by the linear combination of a continuous spectrum of waves with different wave numbers via Fourier analysis. Furthermore, this linear combination must also be a valid solution to the wave equation by the principle of superposition (if ψ_1 and ψ_2 are solutions, $a\psi_1 + b\psi_2$ is also a solution for any constants a and b) which stems from the linearity of the wave equation.

13.2.3 *String Wave*

Let us provide a concrete example of how the one-dimensional wave equation can be obtained in the case of a string wave. A traveling string wave can be produced by perturbing a string under tension. Let the linear mass density of the string when it is relaxed be μ. It is assumed that the tension in the string is large enough that any gravitational effects can be neglected and that each segment of the string produces insignificant longitudinal motion. Every particle on the string moves essentially vertically and the gradient of the string at every point is presumed to remain small.

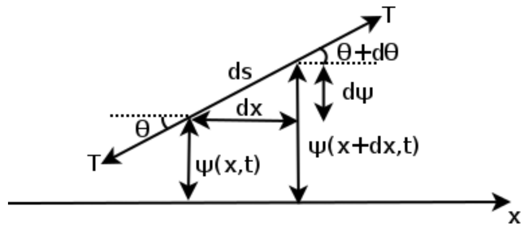


Figure 13.6: String segment

Consider an infinitesimal section of string between equilibrium x-coordinates, x and $x + dx$, at time t (Fig. 13.6). Though the length of this segment is $ds = \sqrt{1 + (\frac{d\psi}{dx})^2} dx$, the mass of this segment is still μdx as the end points are not displaced longitudinally. Instead, the segment is stretched. In order for this infinitesimal section to not accelerate longitudinally, the longitudinal components must be equal. Since the slope of the ends are small, the longitudinal component can be taken to be the actual tension itself (as $\cos\theta$ is second-order and above in θ). Therefore, the tension T must be uniform throughout the string. Applying Newton's second law to this infinitesimal element in the vertical direction and noting that the vertical acceleration of this element is $\frac{\partial^2 \psi}{\partial t^2}$,

$$T \sin(\theta + d\theta) - T \sin\theta = \mu dx \frac{\partial^2 \psi}{\partial t^2}.$$

For small angles, sin can be approximated as tan, and

$$T(\tan(\theta + d\theta) - \tan\theta) = \mu dx \frac{\partial^2 \psi}{\partial t^2}.$$

As the slope of the string is small,

$$\tan\theta \approx \frac{\partial \psi}{\partial x}.$$

Then,

$$T \left(\frac{\partial \psi}{\partial x}\bigg|_{x=x+dx} - \frac{\partial \psi}{\partial x}\bigg|_{x=x} \right) = \mu dx \frac{\partial^2 \psi}{\partial t^2}$$

$$\frac{\frac{\partial \psi}{\partial x}\Big|_{x=x+dx} - \frac{\partial \psi}{\partial x}\Big|_{x=x}}{dx} = \frac{\mu}{T} \frac{\partial^2 \psi}{\partial t^2}.$$

In the limit where $dx \to 0$, from the first principles of calculus,

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{\mu}{T} \frac{\partial^2 \psi}{\partial t^2}, \tag{13.6}$$

which takes the form of the one-dimensional wave equation. It can be seen that a string wave travels at a phase velocity $v = \sqrt{\frac{T}{\mu}}$ where T is the uniform tension and μ is the linear mass density of the string in the relaxed state.

Linear Energy Density

Since a traveling wave transports energy from a source to its surroundings, it is useful to calculate the linear energy density or the energy per unit length of a point on the wave. The energy per unit length of a wave is usually computed instead of the total energy as most waves that we will consider will extend to infinity. The linear energy density of a point on a periodic traveling wave, ε, is generally proportional to the square of the amplitude at that point, i.e.

$$\varepsilon \propto A^2.$$

Let us consider the specific case of a string wave. The kinetic energy of an infinitesimal element of string at coordinate x is

$$dE_{KE} = \frac{1}{2}\mu dx \left(\frac{\partial \psi}{\partial t}\right)^2$$
$$\varepsilon_{KE} = \frac{1}{2}\mu \left(\frac{\partial \psi}{\partial t}\right)^2.$$

The origin of the potential energy is more subtle. When an infinitesimal segment of string is displaced from its equilibrium position, it is stretched by a certain length Δs.

$$\begin{aligned}\Delta s &= ds - dx \\ &= \sqrt{1 + \left(\frac{\partial \psi}{\partial x}\right)^2} dx - dx \\ &\approx \frac{1}{2}\left(\frac{\partial \psi}{\partial x}\right)^2 dx.\end{aligned}$$

The work done by the tension in the string that allows the string to reach this state, dW, is

$$dW = T\Delta s = \frac{1}{2}T\left(\frac{\partial \psi}{\partial x}\right)^2 dx = dU,$$

which is equal to the potential energy stored in this string segment. The potential energy per unit length is then given by

$$\varepsilon_{PE} = \frac{1}{2}T\left(\frac{\partial \psi}{\partial x}\right)^2.$$

Finally, the total energy per unit length is

$$\varepsilon(x,t) = \varepsilon_{KE} + \varepsilon_{PE} = \frac{\mu}{2}\left(\left(\frac{\partial \psi}{\partial t}\right)^2 + \frac{T}{\mu}\left(\frac{\partial \psi}{\partial x}\right)^2\right)$$

$$\varepsilon(x,t) = \frac{\mu}{2}\left(\left(\frac{\partial \psi}{\partial t}\right)^2 + v^2\left(\frac{\partial \psi}{\partial x}\right)^2\right). \tag{13.7}$$

For a traveling wave of the form $\psi = f(x - vt)$ or $\psi = f(x + vt)$, the energy density at a certain point can be rewritten by applying the following relationship

$$\frac{\partial \psi}{\partial t} = \mp v\frac{\partial \psi}{\partial x},$$

$$\varepsilon(x,t) = \mu v^2\left(\frac{\partial \psi}{\partial x}\right)^2 = \mu\left(\frac{\partial \psi}{\partial t}\right)^2, \tag{13.8}$$

which is evidently proportional to the square of the amplitude of the wave at that point. Be wary that Eq. (13.7) is valid for a general wave while Eq. (13.8) only applies to traveling waves. A somewhat counter-intuitive fact is that segments of string that possess the greatest kinetic energy also simultaneously possess the greatest potential energy — a similar statement holds for energy minima.

Power

Since the energy density of a point on a wave is proportional to the square of the amplitude of the wave at that point, the power transmitted through that same point on a periodic traveling wave is also proportional to the squared amplitude of the wave at that point.

$$P \propto A^2.$$

Let us derive an expression for the power transmitted by a string wave at a certain point of coordinate x at time t. Consider a point Q on the string. The transverse force on point Q due to the string segment on the immediate left of Q is $-T\sin\theta = -T\frac{\partial\psi}{\partial x}$. The power transmitted to Q, which is the rate of work done by the left string segment on Q is

$$P = -T\frac{\partial\psi}{\partial x}\cdot\frac{\partial\psi}{\partial t}, \tag{13.9}$$

as $\frac{\partial\psi}{\partial t}$ is the velocity of Q in the transverse direction. Note that we have neglected the longitudinal motion of point Q which is assumed to be non-existent. The above is valid for a general wave, but for a traveling wave of the form $\psi = f(x-vt)$ or $\psi = f(x+vt)$,

$$\frac{\partial\psi}{\partial t} = \mp v\frac{\partial\psi}{\partial x},$$

$$P = \pm Tv\left(\frac{\partial\psi}{\partial x}\right)^2 = \pm\mu v^3\left(\frac{\partial\psi}{\partial x}\right)^2 = \pm\mu v\left(\frac{\partial\psi}{\partial t}\right)^2 = \pm\varepsilon(x,t)v. \tag{13.10}$$

It is evident from the second to last expression that the power transmitted to point Q at time t is proportional to the squared amplitude of the wave at Q. The last expression relating the linear energy density and the speed of the wave makes physical sense. In a time interval dt, an additional length vdt of the wave, which was originally on the left of Q, would have propagated to the right of point Q (we take rightwards to be positive in the x-direction) for a rightward-traveling wave, while carrying an energy density $\varepsilon(x,t)$. Thus, the rate of increase of energy of the string segments on the right of Q is $\varepsilon(x,t)v$. Finally, the above equation can also be used to determine the power delivered by the external agency in maintaining the wave (e.g. by the person perturbing the string).

13.2.4 *Sound Waves*

It is also instructive to analyze sound waves which propagate as varying pressure results in the displacement of gas molecules, which further sparks varying pressure in neighboring regions as the volume occupied by the gas molecules is compressed or expanded due to the initial discrepancy in pressure beyond the equilibrium value. Concretely, consider a tube of gas of uniform density ρ at equilibrium and with cross-sectional area A. We define $\psi(x,t)$ as the longitudinal displacement at time t of the molecules, which were originally located at equilibrium coordinate x, such that the molecules at this instant are at coordinate $x+\psi(x,t)$. Furthermore, let $\psi_p(x,t)$ denote the excess pressure (which may be negative) produced by the molecules at equilibrium coordinate x at time t, beyond the equilibrium pressure p_0.

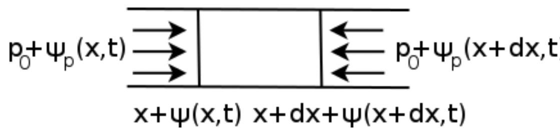


Figure 13.7: Section of gas with ambient pressure p_0

Consider a section of gas with ends at original equilibrium coordinates x and $x + dx$ in Fig. 13.7.

The ends are now located at $x+\psi(x, t)$ and $x+dx+\psi(x+dx)$ respectively. Though the volume of this section is now $V' = A[\psi(x + dx) - \psi(x) + dx]$ instead of the original $V = Adx$, the mass of air in this section is still ρAdx as this gas column still corresponds to the molecules that were originally sandwiched between equilibrium coordinates x and $x + dx$. Observe that there is a change in volume

$$dV = V' - V = A[\psi(x + dx) - \psi(x)],$$

which must be induced by the excess pressure at its ends, ψ_p. For small changes in volume, dV is in fact proportional to the original volume V and the excess pressure ψ_p (for any arbitrary medium).

$$dV = -\kappa V \psi_p,$$

where κ is known as the compressibility of the medium. !times, the bulk modulus $B = \frac{1}{\kappa}$ is also used to describe the change in volume. The negative sign above arises from the fact that increasing the excess pressure reduces the volume of the gas. Shifting V over and substituting the expressions for dV and V,

$$\frac{\psi(x + dx) - \psi(x)}{dx} = -\kappa\psi_p.$$

As $dx \to 0$,

$$\frac{\partial \psi}{\partial x} = -\kappa\psi_p. \tag{13.11}$$

Armed with this relationship, we can now apply Newton's second law to this gas section. The net force is $-[\psi_p(x + dx) - \psi_p(x)]A$ while the mass of gas in this section is ρAdx. Hence,

$$-[\psi_p(x + dx) - \psi_p(x)]A = \rho Adx \frac{\partial^2 \psi}{\partial t^2}.$$

Shifting dx over to the left-hand side,

$$-\frac{\partial \psi_p}{\partial x} = \rho \frac{\partial^2 \psi}{\partial t^2}$$

$$\frac{1}{\kappa}\frac{\partial^2 \psi}{\partial x^2} = \rho \frac{\partial^2 \psi}{\partial t^2}$$

$$\frac{\partial^2 \psi}{\partial x^2} = \rho\kappa \frac{\partial^2 \psi}{\partial t^2}, \tag{13.12}$$

which is the wave equation for displacement. To obtain the wave equation for pressure, we can exploit the commutativity of partial derivatives and take the partial derivative of both sides of Eq. (13.12) with respect to x.

$$\frac{\partial^2}{\partial x^2}\left(\frac{\partial \psi}{\partial x}\right) = \rho\kappa \frac{\partial^2}{\partial t^2}\left(\frac{\partial \psi}{\partial x}\right).$$

Applying Eq. (13.11),

$$\frac{\partial^2 \psi_p}{\partial x^2} = \rho\kappa \frac{\partial^2 \psi_p}{\partial t^2},$$

which is analogous to Eq. (13.12). Therefore, whatever we deduce about ψ or ψ_p is interchangeable. However, as a consequence of Eq. (13.11), note that the wave ψ_p leads ψ by $\frac{\pi}{2}$ radians in the direction of propagation in the case of sinusoidal waves.

It can be seen that the speed of sound is $\sqrt{\frac{1}{\rho\kappa}} = \sqrt{\frac{B}{\rho}}$. This is in fact the general speed of a compressive longitudinal wave in an arbitrary medium — the assumption of the medium being air was not crucial to our derivation. To estimate the compressibility of air, we turn to the adiabatic condition

$$pV^\gamma = c,$$

where p is the pressure of the gas, V is the volume and c is a constant. The process that the gas section undergoes is approximately adiabatic as the time scale of oscillation is much smaller than the time-scale of heat flow. The conductivity σ of an ideal gas is approximately proportional to the product of the mean free path λ (the average distance covered by a molecule between collisions) and the average speed $\langle v \rangle$; $\sigma \propto \lambda\langle v \rangle$. Though the average speed is comparable with the speed of sound, λ is abysmally small (i.e. collisions impede conduction). Furthermore, the adiabatic condition holds especially true for periodic sound waves of long wavelengths as the difference in temperature along the wave is stretched over a long distance. Hence, the rate of conduction is small as the temperature gradient is small — causing the process to be approximately adiabatic. We can take the total derivative of the adiabatic equation to obtain

$$dpV^{\gamma} + \gamma p V^{\gamma-1} dV = 0.$$

Rearranging,

$$dV = -\frac{V}{\gamma p} dp.$$

In most cases, the excess pressure ψ_p is much smaller than p_0. Hence, p in the equation above can simply be taken to be p_0.

$$dV = -\frac{V}{\gamma p_0} dp.$$

Evidently, the compressibility is

$$\kappa = \frac{1}{\gamma p_0}. \tag{13.13}$$

For atmospheric air, $\gamma \approx \frac{7}{5}$ as it is mainly constituted by nitrogen and oxygen which are both diatomic molecules. The speed of a sound wave is thus $\sqrt{\frac{\gamma p_0}{\rho}}$.

Volume Energy Density

We can similarly derive the volume energy density $\varepsilon(x,t)$ of each volume element of a sound wave at a certain time t. The kinetic energy component is evidently

$$\varepsilon_{KE} = \frac{1}{2}\rho \left(\frac{\partial \psi}{\partial t}\right)^2.$$

The potential energy component is again indirect. As the compression or expansion of the gaseous medium is adiabatic, the increase in internal energy in a section of gas of original volume Adx is simply the work done on the gas by the external pressure at its ends due to neighboring sections. Let the total instantaneous pressure on this gas section be p. The increase in internal energy per unit volume in a section of gas from its equilibrium state to a state with excess pressure ψ_p is

$$\frac{-\int p dV}{Adx} = \frac{-\kappa A dx \int_{p_0}^{p_0+\psi_p} p \cdot -dp}{Adx} = \kappa \int_{p_0}^{p_0+\psi_p} p dp = \frac{1}{2}\kappa(2p_0\psi_p + \psi_p^2),$$

where we have applied the relationship $dV = -\kappa V dp = -\kappa A dx dp$ and p_0 is the equilibrium pressure. The term $\kappa p_0 \psi_p$ can be ignored as it averages to zero for oscillatory ψ_p. Then, the potential energy per unit volume stored in a section of gas is

$$\varepsilon_{PE} = \frac{1}{2}\kappa\psi_p^2 = \frac{1}{2\kappa}\left(\frac{\partial\psi}{\partial x}\right)^2.$$

The total volume energy density is thus

$$\varepsilon = \frac{1}{2}\rho\left(\frac{\partial\psi}{\partial t}\right)^2 + \frac{1}{2\kappa}\left(\frac{\partial\psi}{\partial x}\right)^2 = \frac{1}{2}\rho\left[\left(\frac{\partial\psi}{\partial t}\right)^2 + v^2\left(\frac{\partial\psi}{\partial x}\right)^2\right]. \tag{13.14}$$

For traveling waves of the form $\psi = f(x-vt)$ or $f(x+vt)$, $\left(\frac{\partial\psi}{\partial x}\right)^2 = \frac{1}{v^2}\left(\frac{\partial\psi}{\partial t}\right)^2$. Therefore,

$$\varepsilon = \rho\left(\frac{\partial\psi}{\partial t}\right)^2 = \rho v^2\left(\frac{\partial\psi}{\partial x}\right)^2. \tag{13.15}$$

Power

The power transmitted through a surface in space via a planar one-dimensional traveling sound wave can be computed through the rate of work done on an air section by its neighbors on its left. The pressure on the left end of a section is $p_0 + \psi_p$. Therefore, the rate of work done is

$$P = (p_0 + \psi_p)A\frac{\partial\psi}{\partial t}.$$

The $p_0 A\frac{\partial\psi}{\partial t}$ is largely irrelevant here as it averages to zero for periodic ψ. Neglecting this term,

$$P = \psi_p A\frac{\partial\psi}{\partial t} = -\frac{A}{\kappa}\frac{\partial\psi}{\partial x}\cdot\frac{\partial\psi}{\partial t}. \tag{13.16}$$

For traveling waves of the form $\psi = f(x-vt)$ or $f(x+vt)$, $\frac{\partial\psi}{\partial x} = \mp\frac{1}{v}\frac{\partial\psi}{\partial t}$. Then,

$$P = \pm\frac{A}{\kappa v}\left(\frac{\partial\psi}{\partial t}\right)^2 = \pm\rho A v\left(\frac{\partial\psi}{\partial t}\right)^2 = \pm\rho A v^3\left(\frac{\partial\psi}{\partial x}\right)^2, \tag{13.17}$$

where we have applied the relationship $\kappa = \frac{1}{\rho v^2}$. As expected, this result is coherent with $\varepsilon(x,t)Av$.

13.2.5 *Intensity*

Since we are on the topic of power, we slightly digress to an analogous notion for three-dimensional waves. For waves that are inherently three-dimensional (e.g. that emitted by a point light source), it is natural to define a quantity known as the intensity at each point in space. The intensity of a point in space, I, is the average rate at which energy is transported by a wave per unit area across an infinitesimal surface at that point, whose surface is perpendicular to its direction of propagation. In other words, it is the power transmitted per unit area through an infinitesimal surface surrounding a point in space. Since the power of a wave at a spatial position is generally proportional to its squared amplitude at that point in space,

$$I \propto A^2.$$

This relationship, coupled with the fact that energy should be conserved in non-dispersive media (which allows you to indirectly determine intensity), enables us to figure out the amplitude of a symmetric wave without any tedious calculations.

Problem: A point source emits waves isotropically (in the same manner in all directions) whose wave fronts are spherical. Assuming that the total power of the waves does not diminish as they propagate, find the intensity at a point that is a distance r away from the source in terms of the total average power P emitted by the source. Hence, find the amplitude of the wave at that point in terms of A_0, which is the amplitude of the wave at a point that is a distance r_0 from the source.

As the power is evenly distributed about the surface area of the entire sphere with radius r, the intensity at a point on the sphere is

$$I = \frac{P}{4\pi r^2}.$$

This implies that the amplitude is inversely proportional to the radial distance

$$\implies A \propto \frac{1}{r}.$$

Thus,

$$A = \frac{r_0}{r} A_0.$$

It is also not hard to see that for long cylindrical waves (e.g. formed by an infinitely long line of point sources), $A \propto \frac{1}{\sqrt{r}}$ as $I \propto \frac{1}{r}$. Actually, we shall also prove that the amplitude of a spherically symmetric wave decays with $\frac{1}{r}$ directly from the wave equation, in the next section.

13.2.6 *Three-Dimensional Waves*

The three-dimensional wave equation reads

$$\nabla^2 \psi = \frac{1}{v^2}\frac{\partial^2 \psi}{\partial t^2},$$

where ∇^2 is known as the Laplacian operator and v is a constant. Note that the wave function ψ, which represents the displacement at each spatial location at a certain time, must now be a function of all three dimensions.

Plane Waves

The Laplacian in Cartesian coordinates is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

which implies that the wave equation becomes

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} = \frac{1}{v^2}\frac{\partial^2 \psi}{\partial t^2}.$$

As always, we try to look for harmonic solutions as they can be pieced together to constitute a general function by Fourier analysis. To this end, suppose that the solution is separable such that $\psi(x, y, z, t) = X(x)Y(y)Z(z)T(t)$. Substituting this trial solution into the above,

$$X''YZT + XY''ZT + XYZ''T = \frac{1}{v^2}XYZT'',$$

where a prime denotes a differentiation with respect to the function's argument. Dividing the above by $XYZT$,

$$\frac{X''}{X} + \frac{Y''}{Y} + \frac{Z''}{Z} = \frac{1}{v^2}\frac{T''}{T}.$$

Scrutinizing this equation, we find that it consists of terms which are individual functions of x, y, z and t respectively. For example, $\frac{X''}{X}$ is only a function of x while $\frac{T''}{T}$ is strictly a function of t only. For the above to be true for all (x, y, z, t), each of these terms must be a constant! To see why this is so, suppose that we vary t while looking at the displacement of a particular point at equilibrium coordinates (x, y, z) — the right-hand side would have changed (if it were not a constant) while the left-hand side would remain the same. Applying a similar argument to all variables, each term must be a constant. In fact, each constant must be negative, else the corresponding function will blow up at positive or negative infinity (the solution will be exponential). Therefore,

$$\frac{X''}{X} = -k_x^2 \implies X = A_x e^{ik_x x},$$

$$\frac{Y''}{Y} = -k_y^2 \implies Y = A_y e^{ik_y y},$$

$$\frac{Z''}{Z} = -k_z^2 \implies Z = A_z e^{ik_z z},$$

$$\frac{1}{v^2}\frac{T''}{T} = -\frac{\omega^2}{v^2} \implies T = A_t e^{-i\omega t}.$$

Note that we do not include the other exponential solutions whose exponents are negative of the current ones (e.g. $X = A_x' e^{-ik_x x}$ or $T = A_t' e^{i\omega t}$) as the final result obtained from considering the most general solutions can be seen as the superposition of different forms of the solution that we will soon obtain. Concatenating the above expressions together and taking the real component,

$$\psi(x, y, z, t) = A\cos(k_x x + k_y y + k_z z - \omega t + \phi) = A\cos(\boldsymbol{k}\cdot\boldsymbol{r} - \omega t + \phi),$$

where A and ϕ are some constants and

$$\boldsymbol{k} = \begin{pmatrix} k_x \\ k_y \\ k_z \end{pmatrix},$$

which is known as the three-dimensional wave vector. Finally, notice that the separation constants introduced are not independent. They must obey

$$k^2 = k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{v^2}.$$

Now, let us interpret the physical meaning of this wave function. First and foremost, the expression $\psi(x, y, z, t)$ above is known as a plane wave, as all points along a plane perpendicular to $\boldsymbol{k}$ have the same phase at a particular juncture in time (the dot products of the position vectors of all points in such a plane and $\boldsymbol{k}$ are identical). Next, the plane wave is traveling in the direction of $\hat{\boldsymbol{k}}$. This is most obvious when we align our x-axis with $\hat{\boldsymbol{k}}$ such that ψ becomes

$$\psi(x, y, z, t) = A\cos(kx - \omega t + \phi),$$

and we retrieve our one-dimensional sinusoidal wave that is traveling towards the positive x-direction at phase velocity $\frac{\omega}{k}$! Incidentally, this also means that the constant v (positive value) in our three-dimensional wave equation refers to the phase velocity of the wave as $v^2 = \frac{\omega^2}{k^2}$. Finally, it is also evident that $k = \frac{2\pi}{\lambda}$ and $\omega = \frac{2\pi}{T} = 2\pi f$ where λ is the wavelength, T is the period and f is the frequency of the plane wave. However, do not let the components of $\boldsymbol{k}$ fool you into thinking that there exist "wavelengths" in the x, y and z-directions such that we can form a "wavelength vector." Remember that the wavelength of a wave is a scalar quantity and has nothing to do with direction! To convince yourself of this, suppose that we define the "wavelength along a certain direction" as the distance that we have to move along that direction to reach another point that is at the same state of oscillation as the current point that we are at (at the same instance in time). Orient the x-axis along the direction of $\boldsymbol{k}$ such that the "x-component of the wavelength" is λ while the other components tend to infinity. Now, rotate the x and y axes about the z-axis by a certain angle ϕ to produce new x' and y' axes such that the "x' and y'-components of the wavelength" are both finite in general — this contradicts the rotational transformation of vectors (given by multiplying the rotation matrix). Another way to see this is that the magnitude of the two-dimensional vector formed by the "x and y components of the wavelength vector" is not preserved after a rotation about the z-direction — violating a crucial property of vectors.

Spherical Waves

It turns out that the Laplacian in spherical coordinates is

$$\nabla^2 = \frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + \frac{\cos\theta}{r^2 \sin\theta}\frac{\partial}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial \phi^2}.$$

Assuming that our wave function $\psi(r)$ is spherically symmetric such that it is independent of θ and ϕ, the three-dimensional wave equation becomes

$$\frac{\partial^2 \psi}{\partial r^2} + \frac{2}{r}\frac{\partial \psi}{\partial r} = \frac{1}{v^2}\frac{\partial^2 \psi}{\partial t^2}.$$

Introducing a new variable $\psi' = r\psi$,

$$\frac{\partial \psi'}{\partial r} = \psi + r\frac{\partial \psi}{\partial r}$$

$$\implies \frac{\partial^2 \psi'}{\partial r^2} = 2\frac{\partial \psi}{\partial r} + r\frac{\partial^2 \psi}{\partial r^2}$$

$$\frac{\partial^2 \psi'}{\partial t^2} = r\frac{\partial^2 \psi}{\partial t^2},$$

such that the wave equation becomes

$$\frac{\partial^2 \psi'}{\partial r^2} = \frac{1}{v^2}\frac{\partial^2 \psi'}{\partial t^2},$$

which is simply the one-dimensional wave equation that we have analyzed before! Its general sinusoidal solution traveling in the positive r-direction (radially outwards from the origin) is

$$\psi'(r) = A\cos(kr - \omega t + \phi)$$

$$\implies \psi(r) = \frac{A}{r}\cos(kr - \omega t + \phi),$$

where A and ϕ are constants. We have thus proven that the amplitude of a spherically symmetric (isotropic) wave decreases with radial distance r.

13.3 One-Dimensional Waves at a Boundary

13.3.1 *Reflection and Transmission at a Massless Boundary*

Consider two semi-infinite segments of string of linear mass densities μ_1 (in $x < 0$) and μ_2 (in $x > 0$) that are connected at $x = 0$. The tensions in the segments may even differ, for we can connect them via a massless ring that is wrapped around a pole (so that it can absorb the discrepancy in longitudinal tension). Therefore, we let the tensions in the left and right regions be T_1 and T_2 respectively. An interesting question to ask is that if an incident traveling wave propagates from $x = -\infty$ in the positive x-direction, what occurs at the boundary?

Let $\psi_i(x,t) = \psi_i(t - \frac{x}{v_1})$ for $x < 0$ be the incident traveling wave emerging from the left where $v_1 = \sqrt{\frac{T_1}{\mu_1}}$ is the speed of the wave in $x < 0$. Generally, there will be a reflected wave $\psi_r(x,t) = \psi_r(t + \frac{x}{v_1})$ traveling towards the left for $x < 0$ and a transmitted wave $\psi_t(x,t) = \psi_t(t - \frac{x}{v_2})$ traveling towards the right for $x > 0$ where $v_2 = \sqrt{\frac{T_2}{\mu_2}}$. There isn't a "reflected" wave moving towards the left in the region $x > 0$ as the right segment extends to infinity such that no reflection occurs at the right end. The net displacements in the regions $x < 0$ and $x > 0$ are respectively $\psi_i + \psi_r$ and ψ_t.

To determine ψ_r and ψ_t in terms of ψ_i, we have to impose certain boundary conditions at $x = 0$. Firstly, the displacement at $x = 0$ must be continuous at all times as a disjoint string or ring is implausible.

$$\psi_i(0,t) + \psi_r(0,t) = \psi_t(0,t).$$

Furthermore, the transverse components of tensions at the left and right ends of the boundary at $x = 0$ must sum to zero, else the massless entity at $x = 0$ will experience an infinite acceleration. Therefore,

$$T_1\left[\frac{\partial \psi_i}{\partial x}(0,t) + \frac{\partial \psi_r}{\partial x}(0,t)\right] = T_2\frac{\partial \psi_t}{\partial x}(0,t),$$

as $T\frac{\partial \psi}{\partial x}$ is the transverse component of force. As these displacements take the form of traveling waves,

$$\frac{\partial \psi_i}{\partial x}(0,t) = -\frac{1}{v_1}\psi_i'\left(t - \frac{x}{v_1}\right)\bigg|_{x=0} = -\frac{1}{v_1}\frac{\partial \psi_i}{\partial t}(0,t).$$

Note that ($'$) by default denotes differentiation with respect to the function's argument. In the above statement, $\psi_i'(t - \frac{x}{v_1})$ means $\frac{d\psi}{d(t-\frac{x}{v_1})}$. Similarly, $\frac{\partial \psi_r}{\partial x} = \frac{1}{v_1}\frac{\partial \psi_r}{\partial t}$ and $\frac{\partial \psi_t}{\partial x} = -\frac{1}{v_2}\frac{\partial \psi_t}{\partial t}$. Then,

$$-Z_1\frac{\partial \psi_i}{\partial t}(0,t) + Z_1\frac{\partial \psi_r}{\partial t}(0,t) = -Z_2\frac{\partial \psi_t}{\partial t}(0,t),$$

where

$$Z = \frac{T}{v} = \sqrt{T\mu} \tag{13.18}$$

is known as the impedance. Integrating the above and setting the constant of integration to be zero as we assume that there is no displacement before a disturbance reaches $x = 0$, we get

$$Z_1\psi_i(0,t) - Z_1\psi_r(0,t) = Z_2\psi_t(0,t).$$

Combining this requirement and the previous continuity condition,

$$\psi_r(0,t) = \frac{Z_1 - Z_2}{Z_1 + Z_2}\psi_i(0,t),$$

$$\psi_t(0,t) = \frac{2Z_1}{Z_1 + Z_2}\psi_i(0,t).$$

The transmission and reflection coefficients, T and R, are defined as the fractions of the incident amplitude that are transmitted and reflected, respectively. Evidently,

$$R = \frac{Z_1 - Z_2}{Z_1 + Z_2}, \tag{13.19}$$

$$T = \frac{2Z_1}{Z_1 + Z_2}, \tag{13.20}$$

$$T = 1 + R. \tag{13.21}$$

The last condition is enforced by the continuity condition. Before we analyze a few special cases of these coefficients, we proceed to determine $\psi_t(x,t)$ and $\psi_r(x,t)$ from $\psi_t(0,t)$ and $\psi_r(0,t)$. The key is to exploit the fact that the waves are traveling. Therefore,

$$\psi_t(x,t) = \psi_t\left(0, t - \frac{x}{v_2}\right),$$

as the displacement at $x > 0$ at time t would be that at $x = 0$ at time $t - \frac{x}{v_2}$ (the transmitted wave takes $\frac{x}{v_2}$ time to travel from 0 to x). Then,

$$\psi_t(x,t) = \psi_t\left(0, t - \frac{x}{v_2}\right) = T\psi_i\left(0, t - \frac{x}{v_2}\right).$$

Now, even though ψ_i is only valid for $x < 0$, we can imagine the situation where the string is homogeneous with tension T and mass density μ_1 so that we are able to extend its domain to $x > 0$ as well. If ψ_i continued into the $x > 0$ region, its displacement at x-coordinate 0 at time $t - \frac{x}{v_2}$ would be equal to that at x-coordinate $\frac{v_1}{v_2}x$ at time t. Thus,

$$\psi_i\left(0, t - \frac{x}{v_2}\right) = \psi_i\left(\frac{v_1}{v_2}x, t\right).$$

Then,

$$\psi_t(x,t) = T\psi_i\left(\frac{v_1}{v_2}x, t\right).$$

That is, the displacement of ψ_t at coordinate $x > 0$ at a certain instance is T times that of the fictitious incident wave at coordinate $\frac{v_1}{v_2}x$ at the same instance. This implies that the transmitted wave is broadened longitudinally by a factor of $\frac{v_2}{v_1}$ as compared to the incident wave while its amplitude is scaled by a factor of T.

We can apply a similar process to conclude that

$$\psi_r(x,t) = R\psi_i(-x,t),$$

where the domain of ψ_r is $x < 0$. Again, we have extended the domain of ψ_i to the $x > 0$ region as well. The above expression implies that the reflected displacement at coordinate $x < 0$ at a certain instance is that of the imaginary incident wave at $-x > 0$ (i.e. flipped about the point $x = 0$) with an amplitude that is scaled by a factor of R. The "width" of the reflected wave is identical to the incident wave so it is only shrunk transversely.

Special Cases of Impedances

The relative magnitudes of Z_1 and Z_2 determine properties of the reflected and transmitted wave. Before we explore certain special regimes, observe from the expressions

$$R = \frac{Z_1 - Z_2}{Z_1 + Z_2} = -1 + \frac{2Z_1}{Z_1 + Z_2} = 1 - \frac{2Z_2}{Z_1 + Z_2},$$
$$T = \frac{2Z_1}{Z_1 + Z_2} = 2 - \frac{2Z_2}{Z_1 + Z_2},$$

that

$$-1 \leq R \leq 1,$$
$$0 \leq T \leq 2$$

for non-negative impedances.

Case 1: $Z_2 > Z_1$

When the impedance of the second medium exceeds that of the first, $-1 \leq R < 0$ while $0 \leq T < 1$. The negative value of R implies that the reflected wave first undergoes a π-radian phase shift (so that the displacement at $x = 0$ is negated while being scaled) before it is bounced back as shown in Fig. 13.8. This is best visualized by considering a single incident pulse approaching the boundary. A string segment at the $x = 0$ boundary with a positive incident displacement will be pulled down until it attains a negative displacement to form the reflected pulse (note that the net displacement is the superposition of the incident and reflected waves and not only the latter). The reflected pulse is then the incident wave flipped about the x-axis and scaled by a factor of $|R|$. The physical cause of this inversion is the force exerted by the entity at $x = 0$ (e.g. massless ring) on the string segment on its left which wrests it down.

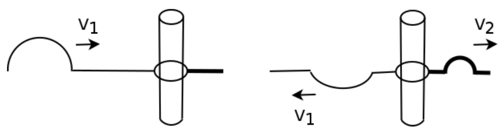


Figure 13.8: Reflected pulse with π-radian phase shift and transmitted pulse

In the limit where $Z_2 \gg Z_1$, $R = -1$ and $T = 0$. That is, all of the incident wave is reflected. A direct consequence of this is that the string segment at $x = 0$ cannot budge, as $\psi_i(0,t) + \psi_r(0,t) = \psi_i(0,t) - \psi_i(0,t) = 0$. Conversely, this also provides an intuitive explanation for the π-radian phase shift as we can coerce the point at the boundary to stay still to achieve infinite impedance.

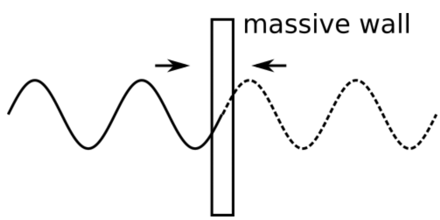


Figure 13.9: Incident wave and imaginary reflected wave

Consider an incident pulse traveling towards the right at speed v_1 and imagine the reflected wave to originate from $x > 0$ and travel towards the left at speed v_1 (Fig. 13.9). The reflected wave is invalid in the $x > 0$ region but the "real" wave emerges in the $x < 0$ region such that the net displacement there is produced by the vector sum of the displacements engendered by the incident and reflected waves. For the node at $x = 0$ to remain stationary at all times, the reflected wave must be a flipped version of the incident one so that their net displacements exactly cancel at $x = 0$ at all times — hence accounting for the π-radian phase shift.

Case 2: $Z_2 < Z_1$

When $Z_2 < Z_1$, $0 < R \leq 1$ and $1 < T \leq 2$. There is no longer a phase shift of π radians upon reflection. Furthermore, in the limit where $Z_1 \gg Z_2$, $R = 1$ and $T = 2$. The displacement at $x = 0$ is in fact twice the displacement that would have been produced by the incident wave alone!

Perhaps, the most intuitive explanation can be obtained from expressing the power transmitted across the boundary (Eq. (13.10)) as $P = Z(\frac{\partial \psi}{\partial t})^2$. Therefore, for small Z_2, there must not be any power transmitted across the boundary (though the transmitted amplitude is larger than normal). Correspondingly, the entity at $x = 0$ cannot exert any transverse force on the string segment on its right and hence also that on its left (for forces to be balanced) — thus maintaining the shape of the left segment by the conservation of energy before bouncing it back. Then, the net displacement at the boundary, which is the vector sum of those due to the incident and reflected waves, must be twice the individual displacement of the incident wave at the boundary.

Case 3: $Z_2 = Z_1$

When $Z_2 = Z_1$, we say that the impedances are matched. Then, $R = 0$ and $T = 1$. That is, there is no reflected wave and the entirety of the incident wave is transmitted. This is intuitive when the string is homogeneous such that the tensions and mass densities are identical on both sides on the boundary, but is less obvious when the two segments are inhomogeneous (e.g. $T_2 = \frac{1}{2}T_1$ and $\mu_2 = 2\mu_1$).

The most important application of this condition lies in the fact that the maximum amount of energy is transmitted across the boundary (as none is reflected). A typical example would be the megaphone whose cross-sectional area is tapered as the impedance of sound propagating across varying cross-sections is inversely proportional to the cross-sectional area. Then, sound can propagate along the megaphone and into the atmosphere with minimal reflection (after sound is channeled and concentrated along the cone).

Physical Meaning of Impedance

In general, for a traveling wave, the force F along the direction of oscillation of a particular point on the wave — exerted by one of its neighbours — is proportional to the velocity of that point, $\frac{\partial\psi}{\partial t}$. The constant of proportionality between them is known as the impedance and is a property of the medium which carries the wave.

$$Z = \frac{F}{\frac{\partial\psi}{\partial t}}. \tag{13.22}$$

Let us first verify that F is indeed proportional to $\frac{\partial\psi}{\partial t}$ for a string. Consider a point on a homogeneous string at coordinate x that is carrying a traveling wave moving towards the right or left, $\psi(x,t) = \psi(x - vt)$ or $\psi(x + vt)$. The transverse component of tension exerted by the segment on the right of x on the left segment is

$$F_{RonL} = T\frac{\partial\psi}{\partial x} = \mp\frac{T}{v}\cdot\frac{\partial\psi}{\partial t} = \mp Z\frac{\partial\psi}{\partial t}, \tag{13.23}$$

where T is the tension in the string. Evidently, F_{RonL} is proportional to $\frac{\partial\psi}{\partial t}$ with the constant of proportionality being $\mp\frac{T}{v}$. The impedance of a string is thus $Z = \frac{T}{v} = \sqrt{T\mu}$ where μ is the mass density of the string. The effect of the right segment is to apply a "damping force" on the left segment with the damping coefficient being the impedance Z. Conversely, the left segment also exerts an equal and opposite force on the right which leads to the delivery of power and the propagation of the wave. The power delivered by the left segment is $F_{LonR} = -F_{RonL}$ multiplied by the velocity of the relevant point $\frac{\partial\psi}{\partial t}$.

$$P = -F_{RonL}\frac{\partial\psi}{\partial t}. \tag{13.24}$$

For rightward and leftward-traveling waves, $F_{RonL} = \mp Z\frac{\partial\psi}{\partial t}$ respectively — implying that

$$P = \pm Z\left(\frac{\partial\psi}{\partial t}\right)^2. \tag{13.25}$$

Moving on, a crucial observation here is that the left segment does not know whether the entity on its right is actually a homogeneous string or a damper with damping constant Z (e.g. a massless plate immersed in water). Therefore, if we replace the right segment with the latter, there should be no impact on the evolution of the left segment as they both feel the same to the left segment. Then, there must be no reflection when the impedance of right entity matches the impedance of the left string segment (as it would think that the medium is homogeneous)! It can then be seen that the impedance is a natural characteristic of a wave, as opposed to other properties of the medium. This is similar to how mass is a good representation of an object's response to a force rather than the object's colour or material, as different objects with the same mass react the same way.

Now, what happens when the impedance is discontinuous at a boundary at $x = 0$? Consider an incident wave ψ_i traveling towards the right from $x = -\infty$ to $x = 0$ along a medium of impedance Z_1 where it encounters a massless damper of impedance Z_2 that is maintained at equilibrium by external means. In general, the viscous force exerted by the damper on the left string segment is not equal to the force required to exactly sustain the incident wave on the left $(-Z_1 \frac{\partial \psi_i}{\partial t}(0,t))$. The excess viscous force then generates a reflected wave ψ_r in the $x \leq 0$ region such that the net displacement is $\psi_i + \psi_r$. The damping force on the left segment required to sustain ψ_r is $Z_1 \frac{\partial \psi_r}{\partial t}(0,t)$. Note the absence of a negative sign as this wave is now traveling towards the left, $\psi_r = \psi_r(x + vt)$. The sum of these forces must be equal to the actual damping force generated by the damper $-Z_2(\frac{\partial \psi_i}{\partial t}(0,t) + \frac{\partial \psi_r}{\partial t}(0,t))$. Therefore,

$$-Z_1 \frac{\partial \psi_i}{\partial t}(0,t) + Z_1 \frac{\partial \psi_r}{\partial t}(0,t) = -Z_2 \left(\frac{\partial \psi_i}{\partial t}(0,t) + \frac{\partial \psi_r}{\partial t}(0,t) \right).$$

Solving for $\frac{\partial \psi_r}{\partial t}(0,t)$,

$$\frac{\partial \psi_r}{\partial t}(0,t) = \frac{Z_1 - Z_2}{Z_1 + Z_2} \frac{\partial \psi_i}{\partial t}(0,t).$$

Integrating with respect to time and imposing the condition that there should be no initial displacement,

$$\psi_r(0,t) = \frac{Z_1 - Z_2}{Z_1 + Z_2} \psi_i(0,t).$$

The rest of the derivation for $\psi_r(x,t)$ follows as above. Finally, we proceed to the case where a medium of impedance Z_2 that extends from $x = 0$ to $x = \infty$ supersedes the damper. The first astute observation to make is that the left segment should respond in the exact same manner as the previous case if the point at $x = 0$ remains at equilibrium (i.e. the transverse components of force on the left and right are continuous) as the left segment cannot determine what is going on at the right side. Thus,

$$\psi_r(0,t) = \frac{Z_1 - Z_2}{Z_1 + Z_2}\psi_i(0,t).$$

Now, the entity at the interface which provides the viscous force on the left segment also provides the driving force to produce a transmitted wave $\psi_t(x,t)$ in the right segment to remain at equilibrium. Again, there is no wave traveling towards the left in region $x \geq 0$ as the only driving force is at $x = 0$. The transmitted wave can be determined by the continuity of displacement.

$$\psi_t(0,t) = \psi_r(0,t) + \psi_i(0,t) = \frac{2Z_1}{Z_1 + Z_2}\psi_i(0,t).$$

Therefore, the reflection and transmission coefficients are $R = \frac{Z_1 - Z_2}{Z_1 + Z_2}$ and $T = \frac{2Z_1}{Z_1+Z_2}$ respectively. The advantage of this formulation is that it is rather general. As long as the component of force along the direction of oscillation at an interface is continuous and the displacement at an interface is continuous (or their analogs), the reflection and transmission coefficients will be given by the above expressions. These boundary conditions happen to hold in most cases — for example, the normal component of velocity $\frac{d\psi}{dt}$ and pressure are continuous across an interface for a sound wave propagating along a constant cross-section. The relevant coefficients for a normally incident sound wave are then given by substituting $Z = \frac{p_0}{v} = \sqrt{\frac{\rho p_0}{\gamma}} = \frac{\rho v}{\gamma}$ for the impedance (the last expression is most common and the γ's can actually be canceled).

13.3.2 *Massive Boundary*

Instead of a purely viscous force, inertia could also be incorporated into the boundary. Then, the boundary condition, in addition to the continuity of displacement, is that the net force on the mass at the boundary, due to the neighboring segments, must produce the correct acceleration as governed by Newton's second law. Consider the following problem.

Problem: Two strings of mass densities μ_1 and μ_2 and tensions T_1 and T_2 are connected via a mediating mass m at $x = 0$ as shown in Fig. 13.10. The mass is constrained such that it can only move in the transverse direction. A continuous and progressive sinusoidal string wave travels towards the right from $x = -\infty$. Determine the reflection and transmission coefficients. Note that there will generally be a phase change in the reflected and transmitted waves, but we are only interested in the amplitudes.

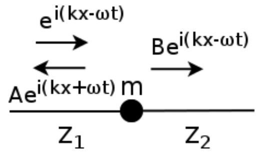


Figure 13.10: Two strings and mediating mass

Let the incident, reflected and transmitted waves be ψ_i, ψ_r and ψ_t respectively. The continuity of displacement at $x = 0$ implies

$$\psi_i(0,t) + \psi_r(0,t) = \psi_t(0,t).$$

The net transverse force on the mass m is $T_2 \frac{\partial \psi_t}{\partial x}(0,t) - T_1 \frac{\partial \psi_i}{\partial x}(0,t) - T_1 \frac{\partial \psi_r}{\partial x}(0,t)$. By Newton's second law,

$$T_2 \frac{\partial \psi_t}{\partial x}(0,t) - T_1 \frac{\partial \psi_i}{\partial x}(0,t) - T_1 \frac{\partial \psi_r}{\partial x}(0,t) = m \frac{\partial^2 \psi}{\partial t^2}(0,t),$$

where ψ on the right-hand side can be taken to be either $\psi_i + \psi_r$ or ψ_t as they are continuous at $x = 0$. We shall choose ψ_t for our purposes. Exploiting the fact that ψ_i, ψ_r and ψ_t refer to traveling waves, the above can be rewritten as

$$-Z_2 \frac{\partial \psi_t}{\partial t}(0,t) + Z_1 \frac{\partial \psi_i}{\partial t}(0,t) - Z_1 \frac{\partial \psi_r}{\partial t}(0,t) = m \frac{\partial^2 \psi_t}{\partial t^2}(0,t).$$

The next step to solving this differential equation is to guess trial solutions. We know that $\psi_i(x,t)$ takes the form of a traveling sinusoidal wave $\cos(kx - \omega t)$ and thus $\psi_i(0,t) = \cos(-\omega t)$. It is then wise to guess sinusoidal solutions for the other displacements at $x = 0$. In fact, it is even more expeditious to express the displacements in terms of complex variables, solve for them and then take the real component to obtain the physical displacement. Since $\psi_i(0,t)$ is sinusoidal, let its complex counterpart be $\tilde{\psi}_i(0,t) = e^{i\omega t}$ (the phase offset does not matter as we are only comparing amplitudes; the exact magnitude does not matter as we are only looking at ratios). Then, we guess exponential solutions to the complex variables

$$\tilde{\psi}_r(0,t) = Ae^{i\omega t},$$
$$\tilde{\psi}_t(0,t) = Be^{i\omega t},$$

where A and B are time-independent constants which are possibly complex. Substituting these expressions into the previous differential equation (after replacing ψ's with the corresponding $\tilde{\psi}$'s),

$$-Z_2 Bi\omega e^{i\omega t} + Z_1 i\omega e^{i\omega t} - Z_1 Ai\omega e^{i\omega t} = m(i\omega)^2 Be^{i\omega t}$$

$$-Z_2 B + Z_1 - Z_1 A = mBi\omega.$$

From the continuity condition,

$$1 + A = B.$$

Solving,

$$A = \frac{Z_1 - Z_2 - im\omega}{Z_1 + Z_2 + im\omega},$$

$$B = \frac{2Z_1}{Z_1 + Z_2 + im\omega}.$$

To obtain the amplitudes of the real waves, observe that for any two complex numbers, $|z_1 z_2| = |z_1||z_2|$. Therefore, the amplitudes of the real waves are $|A|$ and $|B|$ since $|e^{i\omega t}| = 1$. Correspondingly,

$$R = |A| = \frac{\sqrt{(Z_1 - Z_2)^2 + m^2\omega^2}}{\sqrt{(Z_1 + Z_2)^2 + m^2\omega^2}},$$

$$T = |B| = \frac{2Z_1}{\sqrt{(Z_1 + Z_2)^2 + m^2\omega^2}}.$$

Note that these two coefficients no longer need to obey $1 + R = T$ as there are phase differences between the incident, reflected and transmitted waves. These coefficients only represent the amplitudes of the waves and do not account for their phases, which collectively determine the instantaneous displacements.

13.3.3 *Fixed End*

Another common boundary condition comes from restricting the movement of a point on the wave at one end. The general one-dimensional wave problem involves determining the displacement $\psi(x, t)$ of all points at all instances given the initial displacement $\psi(x, 0)$ and velocity $\frac{\partial \psi}{\partial t}(x, 0)$. A wave with a fixed end and zero initial velocity actually provides a convenient avenue for us to solve this general problem, rather than impeding our progress with an additional boundary condition. Let the fixed end be located at $x = 0$ and the region occupied by the wave be located on the right of $x = 0$, extending to infinity (so that the boundary condition on the right can be neglected).

The crux of our approach here is to decompose the general wave into component traveling waves whose superposition satisfy the initial and boundary conditions. Then, the combination of these waves must produce the correct solution as the set-up is completely deterministic. Let us analyze the special case where the initial displacement of the wave is given by

$$\psi(x,0) = f(x),$$

for $x \geq 0$ while the initial velocity $\frac{\partial \psi}{\partial t}(x,0) = 0$ for $x \geq 0$. Obviously, $f(0) = 0$ as the end at $x = 0$ is fixed. Construct the function $g(x)$ such that

$$g(x) = \begin{cases} \frac{f(x)}{2}, & \text{for } x \geq 0 \\ -\frac{f(-x)}{2}, & \text{for } x < 0. \end{cases}$$

That is, $g(x)$ is the half of the original function, $\frac{f(x)}{2}$, in the $x \geq 0$ region but is the negated reflection of $\frac{f(x)}{2}$ about $x = 0$ in the $x < 0$ region. We claim that the solution to the displacement $\psi(x,t)$ is

$$\psi(x,t) = g(x+vt) + g(x-vt),$$

where v is the phase velocity of the wave. Firstly, the above expression is a linear superposition of two functions whose arguments are $(x+vt)$ and $(x-vt)$ — in accordance with D'Alembert's solution. Now, we can check for the initial conditions.

$$\psi(x,0) = 2g(x) = f(x),$$
$$\frac{\partial \psi}{\partial t}(x,0) = vg(x) - vg(x) = 0.$$

These are consistent with the given conditions. Finally, we have to ensure that the end at $x = 0$ remains still at all times.

$$\psi(0,t) = g(vt) + g(-vt) = g(vt) - g(vt) = 0.$$

Since our constructed solution satisfies the wave equation and the initial and boundary conditions, it must be the solution that we are seeking. Now that we have determined this general solution, what does it mean intuitively? It consists of two waves that are vertically shrunk versions (by a factor of half) of the initial displacement that are traveling in opposite directions. This vertical scaling guarantees the validity of the initial displacement while their motions in opposite directions ensure that the wave profile is initially stationary. Next, the waves are fictitiously extended into the $x < 0$ region by reflections so that they are odd functions. This ensures that as one wave travels rightwards and the other leftwards, their displacements at $x = 0$ exactly cancel at all times — thus satisfying the boundary condition. Refer to the following problem for some depictions.

Problem: A string extends from $x = 0$ to $x = \infty$ with a fixed end at $x = 0$. It is initially at rest with an initial displacement the shape of a semi-circle with radius R as shown in the figure below. If the speed of traveling waves[1] on the string is a known value v, prove that at time $t = \frac{R}{2v}$, the displacement of the string at all points is non-negative (above or lying on the horizontal line). Next, draw the shape of the string at time $t = \frac{R}{v}$.

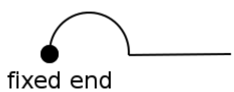


Figure 13.11: Initial displacement

The appropriate $g(x)$ for this wave is

$$g(x) = \begin{cases} \frac{\sqrt{R^2-(x-R)^2}}{2}, & \text{for } x \geq 0 \\ -\frac{\sqrt{R^2-(x+R)^2}}{2}, & \text{for } x < 0 \end{cases}.$$

As depicted below, it consists of a scaled version of the original semi-circle (semi-ellipse) and its reflection.

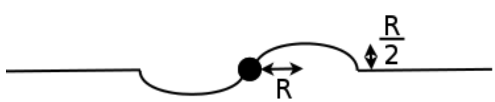


Figure 13.12: $g(x)$ obtained from extending wave into $x < 0$ region

The general solution in this case comprises two of such $g(x)$'s traveling in opposite directions. At $t = \frac{R}{2v}$, each wave would have covered distance $\frac{R}{2}$ and their relative positions are depicted in Fig. 13.13 on the next page (the leftward-traveling wave is represented by the dotted lines).

The net displacement is given by the superposition of the two component waves and can only conceivably be negative in the $0 \leq x < \frac{R}{2}$ region. The

[1]Astute readers may realise that the gradient is definitely not small in this case — an assumption made in the derivation of string waves. However, we can relax this condition and show that the wave equation still holds with T being the constant longitudinal tension (constant so that there is no longitudinal motion). The drawback of this is that the tension can now vary — complicating the energy associated with the wave.

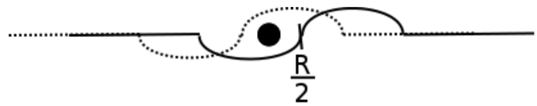


Figure 13.13: Waves at $t = \frac{R}{2v}$

actual net displacement is given by

$$\psi\left(x, \frac{R}{2v}\right) = g\left(x + \frac{R}{2}\right) + g\left(x - \frac{R}{2}\right).$$

For $0 \leq x < \frac{R}{2}$, $x - \frac{R}{2}$ is negative which implies that $g(x - \frac{R}{2}) = -\frac{\sqrt{R^2-(x-\frac{R}{2}+R)^2}}{2} = -\frac{\sqrt{R^2-(x+\frac{R}{2})^2}}{2}$.

$$\psi\left(x, \frac{R}{2v}\right) = \frac{\sqrt{R^2 - \left(x - \frac{R}{2}\right)^2} - \sqrt{R^2 - \left(x + \frac{R}{2}\right)^2}}{2}.$$

This is evidently non-negative for all $0 \leq x < \frac{R}{2}$ as $(x - \frac{R}{2})^2 \leq \frac{R^2}{4}$ while $(x + \frac{R}{2})^2 \geq \frac{R^2}{4}$. At time $t = \frac{R}{v}$, both of the component waves would have travelled a distance R and be located as shown in the figure below.

Figure 13.14: Waves at $t = \frac{R}{v}$

The two waves exactly cancel in the region $0 \leq x \leq R$. However, the leftward-traveling wave yields zero displacement for $x > R$ while the rightward-traveling wave brings along the semi-ellipse. Therefore, the resultant shape of the string is a semi-ellipse of width $2R$ and height $\frac{R}{2}$, centered at $x = 2R$.

13.4 Other Effects

13.4.1 *Polarization*

When light is radiated from filament lamps via spontaneous emissions — a random process where an electron drops from a higher to a lower energy level while emitting a photon in the process — the electric field (and magnetic field) of the resultant electromagnetic wave oscillates, seemingly haphazardly, in many different planes. This is due to the fact that the resultant electromagnetic wave in a certain direction is a superposition of multiple waves which are emitted via random processes and have different planes of

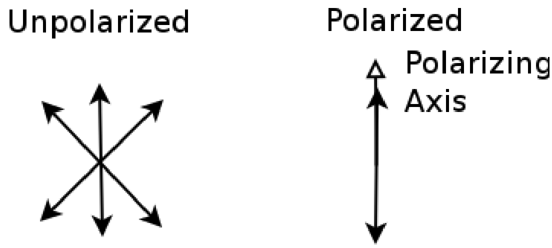


Figure 13.15: Unpolarized and plane-polarized light

oscillation relative to the direction of propagation. Such EM waves with more than one plane of oscillation are known as unpolarized light.

A linear polarizer can be used to restrict the vibrations in a transverse wave, such as an EM wave, to only one direction in the plane normal to the direction of wave propagation. This direction is determined by the direction of the polarizing axis of the polarizer. Only the component of the wave that is parallel to the polarizing axis passes through the polarizer. Note that by this definition, longitudinal waves cannot be polarized as vibrations are always parallel to the direction of propagation in a longitudinal wave. The exact mechanisms of polarization will not be discussed here and we shall just focus on its effects. For an EM wave, the plane of oscillation of its electric field, from the front view, is represented by pairs of arrows which are 180° apart from each other. By convention, only the electric field of an EM wave is depicted, as the magnetic field can be easily obtained by considering the directions of the electric field and wave propagation.

The left diagram in Fig. 13.15 above depicts the oscillations of the electric field of unpolarized light. It shows three planes of oscillations, represented by three pairs of arrows — the two arrows of each pair are 180° apart from each other. After passing through an ideal polarizer with a polarizing axis in the vertical direction as shown above, the oscillation of the electric field of the EM wave that emerges from the polarizer is restricted to the vertical plane alone.

Note that after passing through a polarizer, the magnetic field of the EM wave still exists even though we may not draw it in the diagram. The relationship between the intensities of an initially plane-polarized light before and after passing through a linear polarizer is given by Malus' law. Let the intensities of light before and after passing through the polarizer be I_0 and I respectively. If the angle subtended by the polarizing axis and initially-polarized plane of oscillation is θ, Malus' law states that

$$I = I_0 \cos^2 \theta. \tag{13.26}$$

The squared dependence on $\cos\theta$ stems from two facts — that the amplitude of the electric field is modified by a factor of $\cos\theta$ (component along the polarizing axis), and that the intensity of a EM wave is proportional to the squared amplitude of the electric field.

Problem: Light emerges from a polarizer at intensity I_0. A second polarizer, whose polarizing axis makes a 90° clockwise angle with that of the first, is placed in front of the first polarizer. Determine the orientation of a third polarizer, placed between the first two, that maximizes the transmitted intensity. Determine this maximum intensity.

Let the polarizing axis of the middle polarizer subtend a clockwise angle θ with that of the first. Then, the angle that this makes with the axis of the final polarizer is $90° - \theta$. The transmitted intensity is thus

$$I_0 \cos^2\theta \cos^2(90° - \theta) = I_0 \sin^2\theta \cos^2\theta = \frac{I_0}{4}\sin^2 2\theta,$$

which is maximized when $\theta = 45°$ or $135°$ (note that we do not include $225°$ or $315°$ which represent the same polarizing axes as the two afore). The maximum transmitted intensity is $\frac{I_0}{4}$.

Brewster's Angle

The reflection of light at a boundary between two distinct media with refractive indices n_1 and n_2 can in fact be used to polarize light. Assume that the light starts from medium 1 and enters medium 2 with an angle of incidence θ_i. Let the refracted angle be θ_r. At the special angle of incidence $\theta_i = \theta_b$, such that $\theta_i + \theta_r = 90°$ — known as the Brewster's angle, the oscillation of the electric field of the reflected light is restricted solely to the direction perpendicular to the plane of incidence (the plane formed by the direction of propagation and the normal vector of the interface).
θ_b can be easily computed via Snell's law as

$$\theta_b = \tan^{-1}\frac{n_2}{n_1}. \tag{13.27}$$

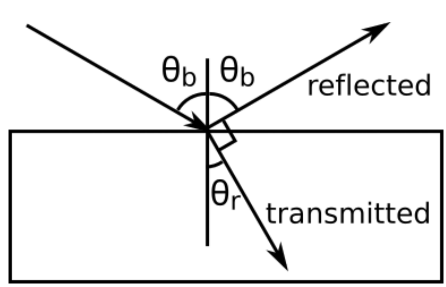


Figure 13.16: Reflection and transmission at Brewster's angle

We shall only provide a qualitative argument of this phenomenon. Firstly, we have to understand the physical origin of reflection and transmission. When an electric field is present in a medium, electric dipoles are induced via various mechanisms (e.g. the nuclei and electron clouds of atoms are slightly separated). Due to the oscillating nature of an EM wave, the dipoles induced by an EM wave oscillate — thus radiating light of their own (as accelerating charges emit radiation). The transmitted and reflected light are then constituted by the radiation of the dipoles (the incident light is ignored as it is diminished due to absorption by the medium). Now, the crux here is that the plane of oscillation of the electric field of the reflected or transmitted light can only be aligned with the direction of oscillation of the dipoles. Furthermore, the plane of oscillation of the electric field must still be perpendicular to the direction of propagation — like any other EM wave.

The direction of oscillation of the dipoles is aligned with the electric field of the refracted wave — which is perpendicular to the direction of the refracted wave. Therefore, the electric field of the reflected wave cannot have a component along the direction of the refracted wave as the dipoles do not oscillate along this direction. Furthermore, the electric field of the reflected wave also cannot have a component along its own direction of propagation. This leaves the only possible direction as that perpendicular to the plane of incidence (perpendicular to the page). Therefore, the electric field of the reflected ray will be polarized along that direction.

13.4.2 *Classical Longitudinal Doppler Effect*

In our daily lives, when a moving siren approaches us or when we are traveling towards a siren, we perceive its pitch to be higher than when it is at rest. Conversely, the pitch of a receding siren is lower. Such perception of pitches is one of our natural instincts that help us to avoid danger and detect swift ambulances. The deviation of the perceived frequency of the waves from its actual frequency of emission from a source is due to the motion of the observer and the source. Let us be quantitative about this.

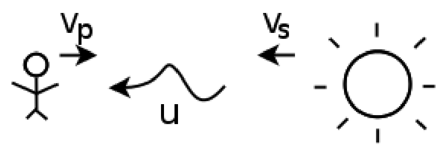


Figure 13.17: Moving source and observer

Consider a source that is traveling at a velocity v_s, towards an observer that is also approaching the source at a velocity v_p in the lab frame (Fig. 13.17). The source emits waves that travel at speed u in the lab frame at a frequency f. What is the frequency of waves reaching the observer (this is the frequency perceived by the observer), assuming that the source and observer are moving directly towards each other? Firstly, we can draw a few wave fronts to get a rough idea of the situation.

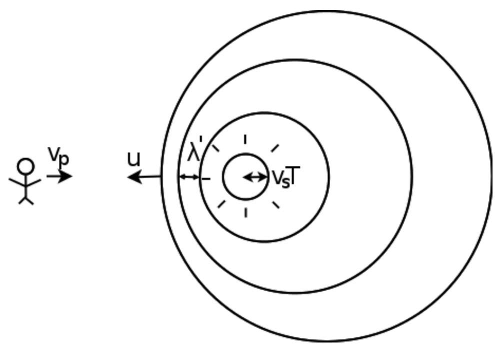


Figure 13.18: Wave fronts emitted by source

The figure above assumes a point wave source for illustrative purposes. Such an assumption is unnecessary but it is easier to visualize the situation with spherical wave fronts. As seen from above, the source is off-center from the spherical wave fronts that have been emitted previously due to the motion of the source during the time period between successive emissions. Let T and λ be the period and wavelength of the waves due to a stationary source and λ' be the wavelength of the waves emitted by the moving source in the lab frame. During a time period T, the wave source would have traveled a distance v_sT. Thus,

$$\lambda' = \lambda - v_sT.$$

Next, the relative velocity between the waves and the observer is $u + v_p$. Thus, the frequency of wave fronts hitting the observer, f', is

$$\begin{aligned} f' &= \frac{u + v_p}{\lambda'} \\ &= \frac{u + v_p}{\frac{\lambda}{T} - v_s} f \end{aligned}$$

$$= \frac{u + v_p}{u - v_s} f$$

$$f' = \frac{u + v_p}{u - v_s} f. \tag{13.28}$$

Take note of the directions of v_p and v_s (towards each other). Otherwise, one can also deduce the sign of the velocities by considering the fact that the observed frequency increases if the observer and the source are approaching each other.

A more in-depth scrutiny of this result would reveal that there must be something wrong with this theory in the case of a wave that does not require a medium for propagation (e.g. light) as the dependencies on the speeds of the observer and source in the lab frame are different! This violates the principle of relativity — a sacrosanct pillar in physics. Suppose that both an observer and a light source were stationary in the lab frame. The observer would receive light of frequency f. Now, the observer goes to sleep and wakes up the next morning to find that the light source is receding from him or her at speed v (perhaps he or she fell asleep on a train). For the sake of simplicity, suppose that we know that it was either the source or the observer which began to move relative to the lab frame (but not both). If this were to be the source, the observer would observe light of frequency $\frac{c}{c+v} f$ where c is the speed of light. In the other case, the observed frequency would be $\frac{c-v}{c} f$ which is different from the previous expression. Therefore, the observer is actually able to discern the entity which actually started moving relative to the lab frame — contravening the supposed uniformity across inertial frames which forbids the determination of the inertial frame that an observer rests in. Ultimately, the observed frequency should only depend on the relative speed between the source and the observer and this loophole in the classical theory is indeed patched by the more accurate theory of special relativity.

Sonic Boom

The Doppler effect only holds for speeds of the source that are lower than that of the wave ($v_s < u$). For $v_s \geq u$, a different phenomenon occurs but the same approach can be applied.

Consider an isotropic point source traveling at speed $v_s > u$ where u is the speed of wave propagation in the lab frame. The wave fronts are spheres that enlarge at speed u which are not concentric due to the movement of the source. The distance between the centers of adjacent wave fronts is $v_s T$ — the distance covered by a source during a period T. Lines can be drawn to connect the different wave fronts to form a cone of half angle θ as shown

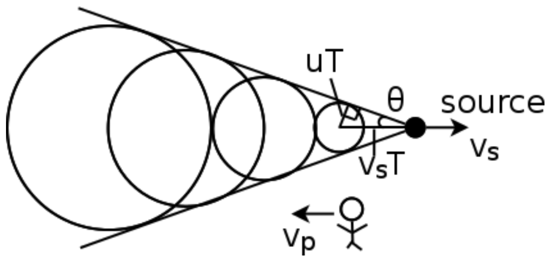


Figure 13.19: Wave fronts emitted by source

in Fig. 13.19. Since the ratio between the distances covered by a sphere and the source is $u : v_s$,

$$\theta = \sin^{-1} \frac{u}{v_s}.$$

Now, observe that the observer only receives a single wave front — this occurs at the juncture where the cone intersects with the observer (i.e. the source has already passed overhead with respect to the observer). Then, the observer perceives a single instantaneous "boom" as the intensity of the wave (most commonly, the sound from a supersonic jet) suddenly increases from zero to that emitted by the source and then decreases to zero shortly afterwards — this stark contrast in intensity leads to a deafening sound.

Problems

1. *Impedance Matching**

A classic example of impedance matching would be the transmission of energy via perfectly elastic collisions between balls. A mass m travels towards a stationary mass M at speed u and undergoes an elastic, head-on collision. Determine the final speeds of m and M. What is the impedance in this case and when is the maximum amount of energy transferred? Finally, devise a way (an infeasible one is fine) to completely transfer energy from m and M via elastic collisions only. You can add more balls.

2. *Linear Expansion**

A rod, with linear mass density μ, Young's modulus Y and initial length l, is attached between two fixed rigid supports. At one temperature, the speed of a longitudinal wave is found to be v_1. When the temperature of the rod is raised by ΔT and the ends of the rod are attached to two new supports, the speed increases to v_2. If the cross-sectional area of the rod is assumed to be a constant A, determine the coefficient of linear expansion α of the rod. Note that the Young's modulus is the stress per unit strain of a section of the rod. The stress σ is the tensile or compressive force that a section of the rod experiences per unit area while the strain ε is the change in the length of a section divided by its original length.

3. *Hanging String**

A string of linear mass density μ and length L is hung vertically from a wall. A mass m is attached to its bottom end, $m \gg \mu L$. If a traveling pulse of a small width, that bulges rightwards, is made at the top of the string and moves towards the bottom end, how does the width of the pulse change qualitatively as it progresses? In what direction does the string move towards? Finally, assuming that the string remains reasonably vertical and stationary (possibly because m is large), determine the time required for the part of the pulse that began at the top, to reach mass m.

4. *Power of Opposite Waves**

Show that the power $P(x,t)$ carried by the superposition of two one-dimensional mechanical waves $\psi_1(x,t) = \psi_1(x-vt)$ and $\psi_2(x,t) = \psi_2(x+vt)$ traveling in opposite directions at speed v in a homogeneous medium of impedance Z is simply the sum of the individual powers. Note that this result is not true in the case of two waves traveling in the same direction.

5. *Reflected Frequency**

A car travels at speed v in the lab frame along horizontal ground and emits waves of speed u in the lab frame at frequency f. If the car is receding from a stationary vertical wall (in a direction normal to the wall), determine the frequency of the reflected waves that are received by the car.

6. *Attenuation***

In reality, waves undergo damping such that their amplitudes decay. The equation describing the waves, obtained from the physical laws (e.g. Newton's laws), then takes the form

$$\frac{\partial^2 \psi}{\partial t^2} + \beta \frac{\partial \psi}{\partial t} = c^2 \frac{\partial^2 \psi}{\partial x^2},$$

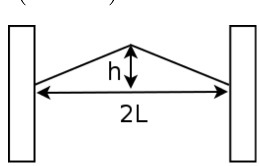
where β and c are positive constants. Determine the solutions to the above equation, that are sinusoidal traveling waves with steady amplitudes (not varying with time). Show that the amplitude is in fact decaying exponentially with distance. Is the phase velocity still independent of wavelength?

7. *Spring Wave***

Determine the speed of a longitudinal traveling wave along a spring of linear mass density μ, spring constant k and length l.

8. *Triangular Wave***

The two ends of a string of mass density μ, relaxed length $2L$ and tension F are fixed to two massive walls separated by a horizontal distance $2L$. The string is initially given a triangular displacement of height h, as shown in the figure below. The string is initially stationary and is then released. Determine the total energy and the period T of the resultant wave. At time $t = \frac{T}{8}$ from the release of the string, draw the shape of the string and label it with the relevant lengths. (APhO)

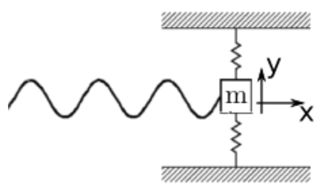


9. *Mass in Middle***

Two strings — each of relaxed length L, tension T and mass density μ — are connected via a mass m in the middle. If the other ends of the strings are fixed to walls a distance L away from m, determine an equation that can be used to solve for the fundamental modes of the system. Now, for $\mu L \gg m$ and $\mu L \ll m$, determine the wavelengths of the fundamental modes.

10. *String Wave and Spring***

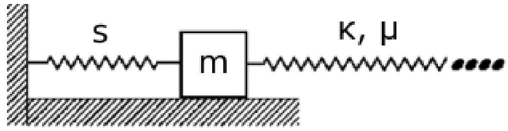
A mass m, located at $x = 0$, is connected to two identical springs with spring constant $\frac{s}{2}$ each and is constrained to move along the y-axis. The springs are both at their rest lengths when m is at the origin. This mass is also tied to a string of tension T. A transverse wave in the string $\psi_i(x,t) = A\sin(\omega t - kx)$ is incident to the mass (refer to the diagram below). The reflected wave is given by $\psi_r = B\sin(\omega t + kx + \phi)$. You may neglect gravity and assume that the gradient of the string is small at all points in the questions below.



(a) Determine the equation of motion of the mass m in the form of a differential equation.
(b) Determine the value of m such that $\phi = 0$ and $B = A$.
(c) This part is not related to (b). Assuming $m = 0$, determine the expressions for $B\sin\phi$ and $B\cos\phi$ in terms of A, s, T and k.
(d) From (c), determine the expressions for B and ϕ. Comment on whether the result makes sense if $s = 0$.

11. *Spring-Mass with Wave***

Referring to the figure on the next page, the left end of a mass m is connected to a wall via a massless spring (spring 1) of spring constant s. The right end of the mass is connected to a very long spring (spring 2). Spring 2 has mass per unit length μ and the value of its spring constant multiplied by its length is κ. Neglect gravity and assume the ground to be frictionless in this problem. Initially, both springs are at their respective equilibrium lengths. At $t = 0$, the mass is given an initial velocity v_0 rightwards which sets up a longitudinal wave in spring 2 that travels rightwards.



(a) Show that the equation of motion of the mass is given by

$$m\frac{d^2X}{dt^2} + \gamma\frac{dX}{dt} + sX = 0$$

where X is the displacement of the mass (the rightwards direction is positive). Determine the constant γ.

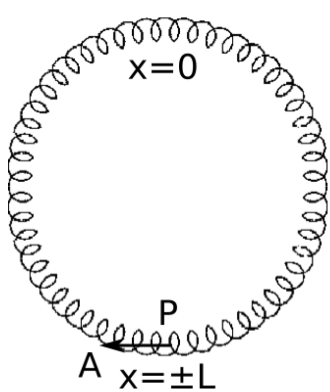
(b) Solve for $X(t)$ in the regime $4ms > \kappa\mu$ using the given initial conditions.

(c) Find the wave function $\psi(x,t)$, where $x = 0$ signifies the equilibrium position of the mass. Note that x here refers to the equilibrium coordinate of a section of spring 2 and $\psi(x,t)$ represents its longitudinal displacement at time t.

(d) The wave function in (c) is only valid for $x < \beta t$. State the constant β and explain why this is so.

12. ***Circular Spring Wave*****

A spring of spring constant k, length $2L$, and mass per unit length μ, is strung along a ring to form a circle, as shown in the figure below. The two ends of the spring are connected at $x = 0$ where the x-coordinate will refer to the position along the ring's perimeter in this problem. Initially, point P, that is located opposite of $x = 0$, is shifted by distance A in the positive x-direction (in this case, to the left) from its equilibrium position, whereas the section at $x = 0$ is held at its original position. Both are released afterwards. We will use $x \in [-L, L]$ in the problems below.



(a) Express the propagation speed of the wave.

(b) Write down all the initial conditions and boundary conditions on the wave function $\psi(x, t)$ that we have for $x \in [-L, L]$. Take the clockwise direction to be positive for displacements.
(c) Plot $\psi(x, 0)$, the initial displacement of each point on the spring from its equilibrium position, for $x \in [-L, L]$.
(d) Determine the period τ of the wave.
(e) Plot the displacement at $t = \frac{\tau}{4}$ (i.e. $\psi(x, \frac{\tau}{4})$ for $x \in [-L, L]$).
(f) Determine the total energy of the wave.

Solutions

1. Impedance Matching*

Let the final velocities of m and M be v_1 and v_2 respectively. By the conservation of momentum,

$$mu = mv_1 + Mv_2.$$

Since the relative velocity between the balls simply reverses during an elastic collision,

$$v_2 - v_1 = u.$$

Solving,

$$v_1 = \frac{m-M}{m+M}u,$$
$$v_2 = \frac{2m}{m+M}u.$$

The impedance in this case is simply the mass of the balls. The maximum energy is transferred when $v_1 = 0$ (i.e. m retains no energy). This occurs when $m = M$. To completely transfer energy from m to M, we can taper mass from m to M by interjecting balls with progressively increasing mass between m and M. We can show that the elastic collision between a ball of mass m' and mass $m' + dm$ completely transfers energy. Let the speed of m' be v. Substituting m' for m and $m' + dm$ for M in the equations above, the energy carried by $m' + dm$ is

$$\begin{aligned} E_t &= \frac{2m'^2(m'+dm)}{(2m'+dm)^2}v^2 \\ &= \frac{m'+dm}{2\left(1+\frac{dm}{2m'}\right)^2}v^2 \\ &\approx \frac{1}{2}(m'+dm)\left(1-\frac{dm}{m'}\right)v^2 \\ &= \frac{1}{2}(m'-dm+dm)v^2 \\ &= \frac{1}{2}m'v^2, \end{aligned}$$

which is the total initial energy carried by m'. For a more general collision between a moving particle m_1 and a stationary particle m_2 characterized by a coefficient of restitution e, the maximum energy transfer to m_2, given fixed e, in fact occurs when their impedances again match (i.e. $m_1 = m_2$). Scrutinizing this fact, this implies that m_1's final velocity is generally non-zero in order to maximize energy transfer — a somewhat counter-intuitive condition. Finally, for a certain range of e, inserting a third mass between m_1 and m_2 can in fact increase the maximum energy transferred to m_2, so we can repeat this operation to further increase the maximum energy transferred. This is despite there being more collisions which dissipate heat. For more details about this general problem, refer to Ref. [3].

2. Linear Expansion*

The wave equation can be derived for an elastic rod in the exact same manner as a one-dimensional sound wave, with Y playing the role of the bulk modulus B. (However, the physics is slightly different. The bulk modulus depends on pressure which is exerted uniformly on the surface of a section of gas, but the Young's modulus depends on the tensile or compressive force in a medium which is one-dimensional. In general, the former describes a change in volume while the latter describes a change in length). Therefore, the speed of a longitudinal wave in the rod is

$$v = \sqrt{\frac{YA}{\mu}}.$$

The discrepancy in speeds stems from the difference in linear mass densities. The length of the rod becomes

$$l' = l(1 + \alpha\Delta T)$$

after its temperature is increased by ΔT. Therefore, its new linear mass density is

$$\mu' = \frac{\mu l}{l'} = \frac{\mu}{1 + \alpha\Delta T}.$$

This implies that

$$v_2 = v_1\sqrt{\frac{\mu}{\mu'}} = v_1\sqrt{1 + \alpha\Delta T} \approx v_1\left(1 + \frac{1}{2}\alpha\Delta T\right).$$

Solving,

$$\alpha = \frac{2\left(\frac{v_2}{v_1} - 1\right)}{\Delta T}.$$

3. Hanging String*

Note that because the tension in the string is now varying, the equation of motion of the string does not yield the wave equation exactly. However, as $\mu L \ll m$, the wave equation holds approximately with the traveling wave having speed $\sqrt{\frac{T}{\mu}}$ along a segment with tension T and mass density μ. Define the origin at the top of the wall and the x-axis to be positive downwards. The tension in the string as a function of x is (for there to be no longitudinal motion)

$$T = [m + \mu(L - x)]g.$$

Therefore, the tension in the string decreases as the pulse travels downwards. As the phase velocity is $v = \sqrt{\frac{T}{\mu}}$, the back of the pulse travels at a greater speed than the front — implying that the pulse decreases in width. Though the amplitude of a transmitted wave is equal to that of the incident wave in this case (the impedance is matched as tension only varies gradually), the width of the pulse decreases — implying that a smaller mass of string bulges towards the right. Since the string is initially vertical, there should be no horizontal force on the string-cum-mass system. Then, the string must rotate rightwards in an attempt to maintain the position of the center of mass. Moving on, the phase velocity of the wave as a function of x is

$$v = \sqrt{\frac{[m + \mu(L - x)]g}{\mu}}.$$

Separating variables and integrating from $x = 0$ (top end) to $x = L$ (bottom end),

$$\int_0^L \frac{1}{\sqrt{\frac{m}{\mu} + L - x}} dx = \int_0^t \sqrt{g} dt$$

$$2\sqrt{\frac{m}{\mu} + L} - 2\sqrt{\frac{m}{\mu}} = \sqrt{g}t.$$

The time taken for the pulse to travel from the top to the bottom end is

$$t = \frac{2\left[\sqrt{\frac{m}{\mu} + L} - \sqrt{\frac{m}{\mu}}\right]}{\sqrt{g}}.$$

4. Power of Opposite Waves*

The power of a wave is generally given by Eq. (13.24) as

$$P = -F_{RonL}\frac{\partial\psi}{\partial t},$$

where F_{RonL} is the force exerted by the right section of the wave on the left section along the direction of oscillation. ψ in this case is $\psi_1 + \psi_2$. Referring to Eq. (13.23), F_{RonL} is $-Z\frac{\partial\psi}{\partial t}$ for a rightward-traveling wave and $Z\frac{\partial\psi}{\partial t}$ for a leftward-traveling wave. The net F_{RonL} is given by the superposition of the forces associated with ψ_1 and ψ_2, i.e.

$$F_{RonL} = -Z\left(\frac{\partial\psi_1}{\partial t} - \frac{\partial\psi_2}{\partial t}\right).$$

Substituting this expression into the above,

$$P = Z\left(\frac{\partial\psi_1}{\partial t} - \frac{\partial\psi_2}{\partial t}\right)\left(\frac{\partial\psi_1}{\partial t} + \frac{\partial\psi_2}{\partial t}\right) = Z\left[\left(\frac{\partial\psi_1}{\partial t}\right)^2 - \left(\frac{\partial\psi_2}{\partial t}\right)^2\right],$$

which is the sum of the two individual powers $Z(\frac{\partial\psi_1}{\partial t})^2$ and $-Z(\frac{\partial\psi_2}{\partial t})^2$ (Eq. (13.25)).

5. Reflected Frequency*

The frequency of waves impinging on the wall is given by the Doppler shift formula as

$$f' = \frac{u}{u+v}f,$$

because the source (the car) is receding. Now, f' is also the frequency of the reflected waves "emitted" by the wall. Therefore, the frequency of reflected waves that is received by the car is

$$f'' = \frac{u-v}{u}f' = \frac{u-v}{u+v}f,$$

as the car now acts as the observer while the stationary wall is the source.

6. Attenuation**

We express a trial sinusoidal solution in terms of complex variables. Substituting

$$\tilde{\psi} = Ae^{i(Kx-\omega t)},$$

where A is a possibly complex constant into the given equation,

$$-\omega^2 Ae^{i(Kx-\omega t)} - i\omega\beta Ae^{i(Kx-\omega t)} = -c^2K^2Ae^{i(Kx-\omega t)}$$

$$c^2K^2 = \omega^2 + i\omega\beta.$$

Now, notice that we can afford for K to be complex but not ω, as a complex component of ω would lead to a real exponential in $\tilde{\psi}$ and hence, a time-varying amplitude. Expressing the previous equation in Euler form,

$$c^2K^2 = \sqrt{\omega^4 + \omega^2\beta^2}e^{i\tan^{-1}\frac{\beta}{\omega}}$$

$$K = \frac{\sqrt[4]{\omega^4 + \omega^2\beta^2}}{c}e^{\frac{i}{2}\tan^{-1}\frac{\beta}{\omega}}.$$

Let the real and complex components of K be k and α respectively. Then,

$$\begin{aligned} k = \text{Re}(K) &= \frac{\sqrt[4]{\omega^4 + \omega^2\beta^2}}{c}\cos\left(\frac{1}{2}\tan^{-1}\frac{\beta}{\omega}\right) \\ &= \frac{\sqrt[4]{\omega^4 + \omega^2\beta^2}}{c}\sqrt{\frac{\cos\left(\tan^{-1}\frac{\beta}{\omega}\right) + 1}{2}} \\ &= \frac{\sqrt[4]{\omega^4 + \omega^2\beta^2}}{c}\sqrt{\frac{\omega}{2\sqrt{\omega^2+\beta^2}} + \frac{1}{2}} \\ &= \frac{1}{c}\sqrt{\frac{\omega^2 + \omega\sqrt{\omega^2+\beta^2}}{2}}. \end{aligned}$$

Similarly, one can show that

$$\alpha = \text{Im}(K) = \frac{1}{c}\sqrt{\frac{\omega\sqrt{\omega^2+\beta^2} - \omega\beta}{2}},$$

but the exact expression is not particularly edifying either. The crucial point is that in general (after expressing A as $Be^{i\phi}$ where B is real),

$$\tilde{\psi} = Be^{-\alpha x}e^{i(kx-\omega t+\phi)}.$$

The real wave is

$$\psi = \mathrm{Re}(\tilde{\psi}) = Be^{-\alpha x}\cos(kx - \omega t + \phi).$$

The phase velocity is

$$v = \frac{\omega}{k} = c\sqrt{\frac{2}{1+\sqrt{1+\frac{\beta^2}{\omega^2}}}},$$

which is a function of ω and thus indirectly, wavelength.

7. Spring Wave**

Define $\psi(x, t)$ as the longitudinal displacement of a point on the spring with equilibrium coordinate x. Consider a section of the spring between equilibrium coordinates x and $x + dx$ with mass μdx. Its ends are currently located at $x + \psi(x, t)$ and $x + dx + \psi(x + dx, t)$. The extension of this section is

$$ds = x + dx + \psi(x + dx, t) - [x + \psi(x, t)] - dx = \psi(x + dx, t) - \psi(x, t).$$

Now, what is the spring constant K of this infinitesimal section? Observe that if we break a spring of spring constant k into two identical pieces connected in series, the extension of each piece is half of that of the original spring but the tension must remain the same — implying that the spring constant doubles. In general, the spring constant of a section multiplied by the length of the section is a constant. Therefore,

$$K dx = kl.$$

Now, apply Newton's second law to this spring section. The force on this section due to its right neighbor is $K[\psi(x + dx, t) - \psi(x, t)]$ while that due to its left neighbor is $-K[\psi(x, t) - \psi(x - dx, t)]$. By Newton's second law,

$$\mu dx \frac{\partial^2 \psi}{\partial t^2}(x, t) = K[\psi(x + dx, t) - \psi(x, t)] - K[\psi(x, t) - \psi(x - dx, t)],$$

where $\psi(x, t)$ can be taken to be the displacement of the center of mass of this section as dx is small. Substituting $K = \frac{kl}{dx}$ and applying the first principles of calculus,

$$\mu dx \frac{\partial^2 \psi}{\partial t^2} = kl \left(\frac{\partial \psi}{\partial x}\Big|_{x=x+dx} - \frac{\partial \psi}{\partial x}\Big|_{x=x} \right).$$

Shifting μdx over to the right-hand side and applying the first principles of calculus again,

$$\frac{\partial^2 \psi}{\partial t^2} = \frac{kl}{\mu} \frac{\partial^2 \psi}{\partial x^2}.$$

The phase velocity of the spring wave is evidently $\sqrt{\frac{kl}{\mu}}$.

8. Triangular Wave**

The total energy carried by the wave is the initial potential energy stored, as energy is conserved. Define the origin to be at the left end of the string. The equations of the left and right segments are $y = \frac{h}{L}x$ and $y = h - \frac{h}{L}x$ respectively. The magnitude of the gradients is $\frac{h}{L}$. Therefore, the stored potential energy is

$$\begin{aligned} E &= \int_0^{2L} \frac{1}{2} F \left(\frac{\partial \psi}{\partial x} \right)^2 dx \\ &= \int_0^{2L} \frac{Fh^2}{2L^2} dx \\ &= \frac{Fh^2}{L}. \end{aligned}$$

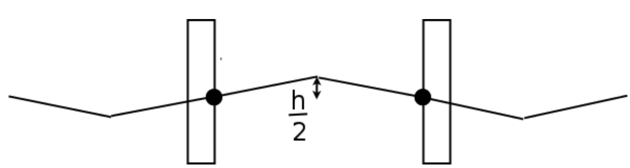


Figure 13.20: Shape of component wave

To determine the period of the resultant wave, observe that the resultant wave can be decomposed into the superposition of two smaller triangular waves (scaled down vertically by a factor of $\frac{1}{2}$ and extended beyond the walls, as shown in Fig. 13.20) traveling in opposite directions at speed $v = \sqrt{\frac{F}{\mu}}$.

The actual wave is composed of two of the above "sub-waves" traveling in opposite directions. After the two waves have "covered" $4L$ distance each (this is the wavelength), the string returns to its original state. Therefore, the period is

$$T = \frac{4L}{v} = \frac{4L\sqrt{\mu}}{\sqrt{F}}.$$

At time $t = \frac{T}{8}$, the two waves would have traveled distance $\frac{L}{2}$ in opposite directions — their positions at this juncture are depicted below.

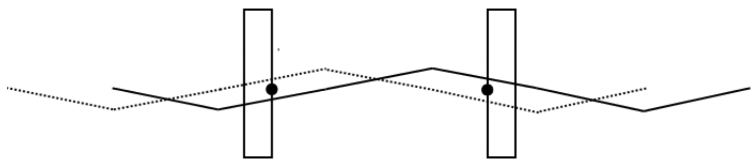


Figure 13.21: Component waves at $t = \frac{T}{8}$

Their superposition yields the following trapezoid.

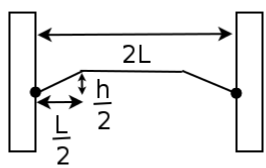


Figure 13.22: Superposition

Note that the interval where the two gradients of the component waves are opposite consists of a flat plateau. Also, each component wave has a maximum height $\frac{h}{2}$.

9. Mass in Middle**

Set the origin at the mass and the coordinates of the fixed ends to be $-L$ and L respectively. Let the waves on the left and right of m be represented by ψ_L and ψ_R. Since the end at $-L$ is fixed, the expression for ψ_L must take the form

$$\psi_L = A\sin(k(x+L)+n\pi)\sin(\omega t+\phi)$$

where n is an integer and ϕ is a constant phase offset (see next chapter for the equation for standing waves or the superposition of two identical waves traveling in opposite directions). Similarly, as the end at L is fixed, the expression for ψ_R must take the form

$$\psi_R = A\sin(k(L-x)+n\pi)\sin(\omega t+\phi).$$

The amplitude of ψ_R must be equal to that of ψ_L to ensure the possibility of continuity at $x=0$ at all times. Furthermore, the time-dependences of ψ_R and ψ_L must be identical to ensure continuity. Actually, the $n\pi$ in ψ_R can be replaced by the addition of any even multiple of π to $n\pi$ to enforce continuity, but this yields the same expression. Now, the net force on the mass m must be coherent with its acceleration as given by Newton's second law.

$$T\frac{\partial\psi_R}{\partial x}(0,t) - T\frac{\partial\psi_L}{\partial x}(0,t) = m\frac{\partial^2\psi}{\partial t^2}(0,t),$$

where ψ on the right-hand side can be taken to be either ψ_L or ψ_R as they are continuous at $x=0$. Substituting the expressions for ψ_L and ψ_R and choosing $\psi=\psi_L$,

$$-2TkA\cos(kL+n\pi)\sin(\omega t+\phi) = -m\omega^2 A\sin(kL+n\pi)\sin(\omega t+\phi).$$

Applying $\omega^2 = v^2k^2 = \frac{Tk^2}{\mu}$ and rearranging,

$$\tan(kL+n\pi)kL = \frac{2\mu L}{m}.$$

Now, we can remove $n\pi$ from the argument of tan since it has a period of π. Note that we added the $n\pi$ initially as a precaution — if the final result involved a cos, it would matter if n were even or odd.

$$\tan(kL)kL = \frac{2\mu L}{m}.$$

When $\mu L \gg m$, $\tan(kL)$ must tend to infinity. Then,

$$kL = \frac{\pi}{2} + p\pi,$$

where p is a non-negative integer. Using $k=\frac{2\pi}{\lambda}$, the possible wavelengths are

$$\lambda = \frac{4L}{2p+1},$$

which makes intuitive sense as these correspond to the cases where an antinode is located at m (so that it oscillates wildly). When $\mu L \ll m$, the situation is rather special. When kL itself is reasonably small, $\tan kL \approx kL$. Then,

$$k^2L^2 = \frac{2\mu L}{m},$$

$$\lambda = \pi\sqrt{\frac{2mL}{\mu}}.$$

This is the longest wavelength of the system. For other larger values of kL, $\tan(kL)$ itself must tend to zero. That is,

$$kL = p\pi$$

for some positive integer p. Then,

$$\lambda = \frac{2L}{p}.$$

This also makes sense as such values of λ correspond to the modes where a node is located at m (as it is too ponderous to move).

10. String Wave and Spring**

Let the vertical displacement of m be $\psi(0,t)$. The vertical forces on m are the spring force $-s\psi(0,t)$ and the vertical component of the tension on m due to the left string segment, $-T\frac{\partial\psi}{\partial x}(0,t)$ (note that we are using $\tan\theta$ where θ is the angle subtended by the string segment at $x = 0$ and the horizontal, instead of the more precise $\sin\theta$, as θ is small when the gradient of the string is small). The equation of motion of m is thus

$$m\frac{\partial^2\psi}{\partial t^2}(0,t) = -s\psi(0,t) - T\frac{\partial\psi}{\partial x}(0,t).$$

For part (b), the solution for $\psi(x,t)$ is

$$\psi(x,t) = \psi_i(x,t) + \psi_r(x,t) = A[\sin(\omega t - kx) + \sin(\omega t + kx)],$$

when $B = A$ and $\phi = 0$.

$$\implies \frac{\partial\psi}{\partial x}(x,t) = -kA\sin(\omega t - kx) + kA\sin(\omega t + kx),$$

$$\frac{\partial\psi}{\partial x}(0,t) = 0.$$

Furthermore,

$$\frac{\partial^2\psi}{\partial t^2}(0,t) = -\omega^2\psi(0,t).$$

Therefore, for this solution $\psi(x,t)$ to satisfy the boundary condition at m,

$$(s - m\omega^2)\psi(0,t) = 0$$
$$\implies m = \frac{s}{\omega^2}.$$

For part (c), the boundary condition at $x = 0$ becomes

$$s\psi(0,t) = -T\frac{\partial \psi}{\partial x}(0,t)$$

when $m = 0$. Substituting $\psi(x,t) = A\sin(\omega t - kx) + B\sin(\omega t + kx + \phi)$, the above becomes

$$s[A\sin(\omega t) + B\sin(\omega t + \phi)] = T[Ak\cos(\omega t) - Bk\cos(\omega t + \phi)].$$

To solve for $B\sin\phi$ and $B\cos\phi$, we can substitute $\omega t = 0$ and $\omega t = \frac{\pi}{2}$ to obtain

$$B(s\sin\phi + Tk\cos\phi) = ATk,$$
$$B(s\cos\phi - Tk\sin\phi) = -sA.$$

Solving the above simultaneously,

$$B\sin\phi = \frac{2ATks}{s^2 + T^2k^2},$$
$$B\cos\phi = \frac{(T^2k^2 - s^2)A}{s^2 + T^2k^2}.$$

Solving for B and ϕ,

$$B = A,$$
$$\phi = \cos^{-1}\frac{T^2k^2 - s^2}{s^2 + T^2k^2}.$$

When $s = 0$, $B = A$ and $\phi = 0$. This makes sense as the situation, when the springs are absent, reduces to that of the reflection of a wave that approaches a medium with zero impedance at $x = 0$ for which the reflection coefficient is $R = 1$ (bounce the incident wave back without a vertical flip).

11. Spring-Mass with Wave**

The forces that m experiences are due to the two springs. When m is at an instantaneous displacement X from its equilibrium position rightwards, the force exerted on it by spring 1 is evidently $-sX$. The force due to spring

2 can be computed as follows. Consider the section of spring 2 whose ends were originally between equilibrium x-coordinates 0 and dx. After the spring wave has propagated, its ends are located at coordinates $0+\psi(0,t)$ and $dx+\psi(dx,t)$. Evidently, this section has been stretched by a distance $\psi(dx,t)-\psi(0,t)$. Referring to the solution to Problem 7, the spring constant of this section is $\frac{\kappa}{dx}$. Therefore, the tension at its ends and the force that it exerts on m are

$$F = \kappa \frac{\psi(dx,t)-\psi(0,t)}{dx} = \kappa \frac{\partial \psi}{\partial x}(0,t),$$

by Hooke's law. For a general rightward-traveling wave $\psi(x,t) = f(x-vt)$ where v is the phase velocity of the wave, we have

$$\frac{\partial \psi}{\partial x}(x,t) = -\frac{1}{v}\frac{\partial \psi}{\partial t}(x,t)$$
$$\implies F = -\frac{\kappa}{v}\frac{\partial \psi}{\partial t}(0,t).$$

Substituting $v = \sqrt{\frac{\kappa}{\mu}}$ for a spring wave (see Problem 7) and $\psi(0,t) = X(t)$ (boundary condition at $x=0$),

$$F = -\sqrt{\kappa\mu}\frac{dX}{dt}.$$

Applying Newton's second law to m,

$$m\frac{d^2X}{dt^2} = -sX - \sqrt{\kappa\mu}\frac{dX}{dt}$$
$$m\frac{d^2X}{dt^2} + \sqrt{\kappa\mu}\frac{dX}{dt} + sX = 0,$$

which shows that $\gamma = \sqrt{\kappa\mu}$. This equation simply describes a damped oscillation. Its general solution, in the regime $4ms > \kappa\mu$ which corresponds to light damping, was derived in the chapter on oscillations as

$$X = e^{-\frac{\sqrt{\kappa\mu}}{2}t} c \sin\left(\sqrt{\frac{s}{m} - \frac{\kappa\mu}{4}}t + \phi\right),$$

where c and ϕ are constants determined by initial conditions. Substituting $X = 0$ at $t = 0$, we obtain $\phi = 0$. Substituting this into the above and differentiating it with respect to time t,

$$\frac{dX}{dt} = -\frac{\sqrt{\kappa\mu}}{2}e^{-\frac{\sqrt{\kappa\mu}}{2}t}c\sin\left(\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}t\right)$$
$$+e^{-\frac{\sqrt{\kappa\mu}}{2}t}\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}c\cos\left(\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}t\right).$$

Imposing the initial condition $\frac{dX}{dt} = v_0$ at $t = 0$,

$$c = \frac{v_0}{\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}}$$
$$\implies X = e^{-\frac{\sqrt{\kappa\mu}}{2}t}\frac{v_0}{\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}}\sin\left(\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}t\right),$$

$\psi(x,t)$ can be determined by exploiting the fact that the wave is traveling rightwards at velocity $v = \sqrt{\frac{\kappa}{\mu}}$.

$$\psi(x,t) = \psi\left(0, t-\frac{x}{v}\right) = \psi\left(0, t-\sqrt{\frac{\mu}{\kappa}}x\right),$$
$$\psi(x,t) = e^{-\frac{\sqrt{\kappa\mu}}{2}\left(t-\sqrt{\frac{\mu}{\kappa}}x\right)}\frac{v_0}{\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}}\sin\left[\sqrt{\frac{s}{m}-\frac{\kappa\mu}{4}}\left(t-\sqrt{\frac{\mu}{\kappa}}x\right)\right].$$

This expression is only valid for $x \leq vt = \sqrt{\frac{\kappa}{\mu}}t$ as the traveling wave evidently has yet to reach $x > vt$.

12. Circular Spring Wave**

The phase velocity v of a spring wave was derived in Problem 7 as

$$v = \sqrt{\frac{2kL}{\mu}}.$$

Initially, we can divide the spring into two sections $x \in [-L, 0]$ and $x \in [0, L]$ which are each stretched or compressed uniformly for a total distance A each.

Therefore,

$$\psi(x,0) = -\frac{Ax}{L}, \qquad (-L \leq x \leq 0)$$

$$\psi(x,0) = \frac{Ax}{L}. \qquad (0 \leq x \leq L)$$

Furthermore, all sections of the spring are initially stationary, hence

$$\frac{\partial \psi}{\partial t}(x,0) = 0. \qquad (-L \leq x \leq L)$$

The boundary conditions are that the displacement and velocity must be continuous at $x = -L$ and $x = L$ as they correspond to the same physical section:

$$\psi(-L,t) = \psi(L,t),$$

$$\frac{\partial \psi}{\partial t}(-L,t) = \frac{\partial \psi}{\partial t}(L,t).$$

$\psi(x,0)$ is plotted below.

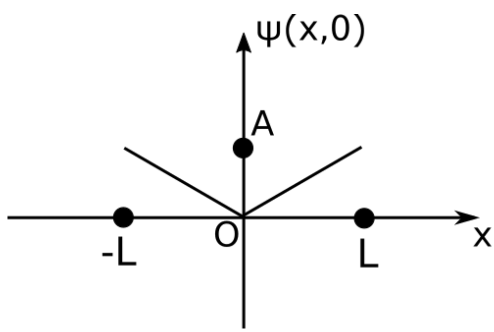


Figure 13.23: $\psi(x,0)$ against x

For part c), the general solution to the wave function can be seen as the superposition of two component waves which are traveling at phase velocity v in opposite directions (positive and negative x-directions). The shape of one component wave ψ_c at $t = 0$ is depicted in Fig. 13.24 (it extends to positive and negative infinity but the physical wave is only valid in the regime $-L \leq x \leq L$).

One can easily check that this superposition yields the correct initial and boundary conditions, but let us explain the intuition behind such a construction. Firstly, the component wave should be a version of the original wave scaled down by a factor of half in the region $x \in [-L, L]$ initially, such that two component waves traveling in opposite directions will satisfy

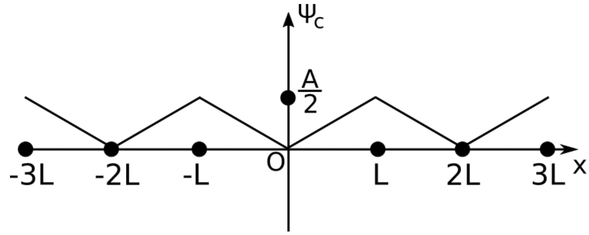


Figure 13.24: Shape of component wave

the initial conditions on displacement and velocity. To develop the shape of the component wave beyond this region, we notice that if the component wave traveled rightwards for a distance d, the part of the wave originally between $[L-d, L]$ would have vanished. This part that disappears should now reappear on the left, in the region $[-L, -L+d]$, because the real wave is cyclical such that the parts that traveled rightwards beyond $x = L$ should be continued at $x = -L$. A similar statement holds for a component wave traveling towards the left. Finally, observe that with this construction, the displacement of each spring section repeats itself after the component waves have traveled a distance $2L$ each. Therefore, the period of the original wave is

$$\tau = \frac{2L}{v} = \sqrt{\frac{2L\mu}{k}}.$$

At $t = \frac{\tau}{4}$, each component wave would have traveled a distance $\frac{L}{2}$. Their shapes at this juncture are thus illustrated in Fig. 13.25 — their superposition yields a horizontal line with displacement $\frac{A}{2}$.

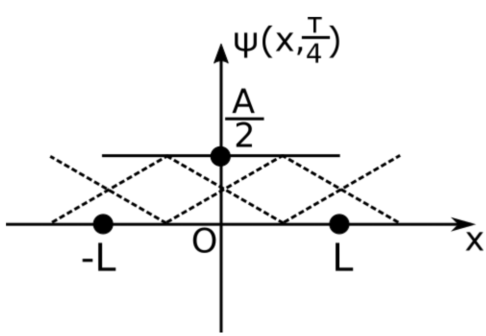


Figure 13.25: Shape of component waves at $t = \frac{\tau}{4}$ and superposition

Finally, as the total energy of the wave is conserved, the total energy is simply the initial potential energy — this is the potential energy of two springs with spring constant $2k$ ($2k$ such that they produce the original spring constant k when they are connected in series) which are each uniformly compressed and extended by a total distance A. Therefore,

$$E = 2 \cdot \frac{1}{2} \cdot 2kA^2 = 2kA^2.$$

This page intentionally left blank

Chapter 14

Interference

The previous chapter discussed the properties of a single traveling wave, with a focus on sinusoidal wave forms. In general, a wave form can take the form of an arbitrary shape. However, it can always be approximated by the linear, algebraic summation of multiple sinusoidal waves that are overlapping via Fourier analysis. Hence, understanding how a myriad overlapping sinusoidal waves, of the same type, can be "added" is pivotal in the study of waves and shall be the crux of this chapter. Such coinciding waves are known to be interfering with one another or undergoing interference.

14.1 The Principle of Superposition

To study how waves of the same type "interact" with one another, consider the following example. Person A wiggles the left end of a uniform rope and creates a wave form traveling towards the right while person B wiggles the right end of the rope and creates a wave form traveling to the left. Both people constantly pull the rope such that the rope is taut with a constant tension along its segments.

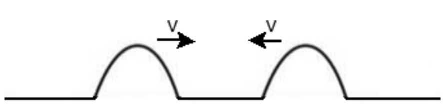


Figure 14.1: Approaching waves

What will be the resultant displacement of the portion of rope at which the waves overlap or partially overlap at a particular instant? It turns out that the resultant displacement at a point at a particular instant, due to multiple overlapping waves of the same nature, is simply the vector sum of the displacements that each individual wave would have caused at that point in space and time. This is known as the principle of superposition.

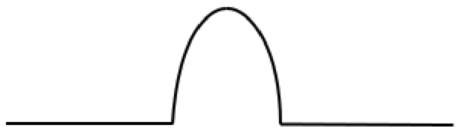


Figure 14.2: Intersecting waves

After the two wave forms pass each other, they retain their "shape" and continue their motion as if they have never met.

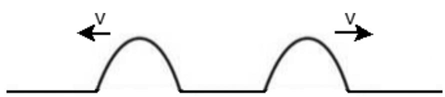


Figure 14.3: Departing waves

This property of waves is rather unique as they can pass through each other and be at the same place at the same time — a stark contrast with the characteristics of particles. However, this is also to be expected as a traveling wave is not a physical entity that exists and translates in space. It is merely a coordinated and coupled movement of points, similar to a Mexican wave performed by ardent sports fans. Using the same analogy, if the peaks of two identical human waves that travel in opposite directions coincide in the middle of the crowd, the fans at the center will simply jump twice the normal height with twice the enthusiasm.

The principle of superposition stems from the linearity of the wave equation

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{v^2}\frac{\partial^2 \psi}{\partial t^2}.$$

If $\psi_1(x,t)$ and $\psi_2(x,t)$ are individual solutions to the wave equation, $\psi_1(x,t) + \psi_2(x,t)$ is also a valid solution. Referring to the previous situation, $\psi_1(x,t)$ and $\psi_2(x,t)$ may be produced individually by persons A and B respectively. If they were to perturb the rope concurrently (at the same location and time), the displacement would be $\psi_1(x,t)+\psi_2(x,t)$ at any point in space and time.

Finally, multiple waves of the same type (e.g. two sound waves) that overlap at a particular point in space and time are said to be interfering.

The resultant displacement[1] at any point in space and time is given by the principle of superposition.

14.1.1 *Constructive and Destructive Interferences*

Consider a snapshot of two sinusoidal waves, P and Q, that travel in the positive x-direction, have the same wavelength λ and frequency f but are not necessarily "perfectly in sync" at a particular time. That is, we are considering how the displacement of various points of the waves look like at a particular instant.

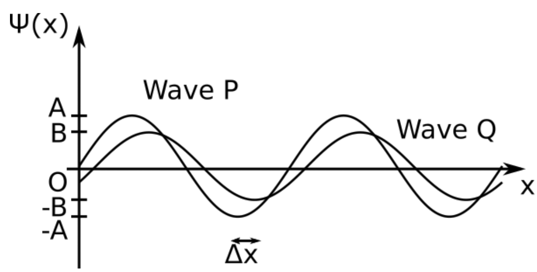


Figure 14.4: Overlapping waves

As observed from the diagram above, a point on wave P at a particular x-coordinate has a phase (state of oscillation) that is different from that of the corresponding point on wave Q. This is evident from the disparity in the locations of the peaks and troughs — when a point at a certain x-coordinate on Q has reached the maximum, the point with the same x-coordinate on P has yet to reach it. This implies a **phase difference**, $\Delta\phi$ between every pair of corresponding points on the waves that are at the same x-coordinate. Mathematically, recall that the general equation of a sinusoidal wave traveling in the positive x-direction is

$$\psi(x,t) = A\sin(kx - \omega t + \phi), \tag{14.1}$$

where $k = \frac{2\pi}{\lambda}$ is the wave number, ω is the angular frequency $\omega = 2\pi f$ and ϕ is an arbitrary constant phase offset. Since we are only considering the traveling wave at a certain instance, we substitute a constant for t to remove the dependence of the above equation on t. Let the individual displacement

[1] In the case of light, there is no medium that carries it and hence no displacement of physical particles. It is actually the electric and magnetic fields that oscillate but we shall still refer to them as "displacements" for the sake of convenience.

of a particle at coordinate x due to wave P be

$$\psi_P(x) = A\sin(kx). \tag{14.2}$$

We can always choose a time t such that the form above holds — this does not affect our calculation of the phase difference as the two waves have the same wavelength and travel at the same speed ($v = f\lambda$) such that the distance Δx between consecutive peaks and troughs is preserved (the importance of this will be revealed soon). Next, the individual displacement of a particle at coordinate x due to wave Q is

$$\psi_Q(x) = B\sin(kx - \Delta\phi) \tag{14.3}$$

for some constant $\Delta\phi$, which is known as the phase difference. Wave Q is said to lead wave P in this case (this statement is made while taking into account the direction of travel). $\Delta\phi$ can be expressed in terms of Δx by observing that the phases of the points on wave Q at this instance are identical to those of the points on wave P after it has travelled an additional distance Δx. Thus,

$$\begin{aligned}\psi_Q(x) &= B\sin(k(x - \Delta x)),\\ \frac{2\pi}{\lambda}(x - \Delta x) &= \frac{2\pi}{\lambda}x - \Delta\phi,\\ \frac{\Delta\phi}{2\pi} &= \frac{\Delta x}{\lambda}.\end{aligned} \tag{14.4}$$

Equation (14.4) describes the phase difference $\Delta\phi$ between one wave that leads another. The reason behind the head start is irrelevant — perhaps, the source of Q was located at a larger x-coordinate than that of P or perhaps, the sources of P and Q were at the same position but the source of Q emitted waves that already had a constant phase difference with respect to those produced by the source of P. Incidentally, there is another disparate effect that is described by Eq. (14.4). Consider two points on the **same** wave at two different x-coordinates x and $x + \Delta x$. By Eq. (14.1), the difference in the phase angle $\Delta\phi$ of these two points is also given by Eq. (14.4). We shall encapsulate all of these factors into the single equation above.

Moving on, in this chapter, all wave sources are assumed by default to be **coherent** — a condition which means that the phase difference between the waves emitted by the sources remains constant over time. In order for two sources to be coherent, they must have the same frequency. The coherency of waves P and Q enabled us to consider an arbitrary time at which wave P took the form of Eq. (14.2). Furthermore, we shall now see that coherency ascribes meaning to the phase difference.

There are two special cases of $\Delta\phi$ to consider. When the phase difference is an even multiple of π (i.e. $\Delta\phi = 2n\pi$, $n \in \mathbb{Z}$), the two waves are "fully aligned" (Fig. 14.5).

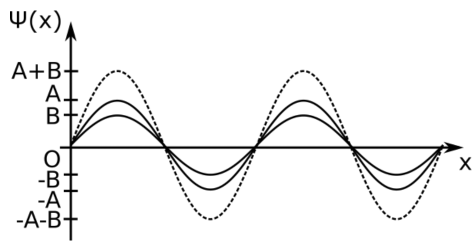


Figure 14.5: Constructive interference

Then, waves P and Q are said to be in phase as each pair of corresponding points at the same x-coordinate literally has the same phase. When corresponding points on different waves have the same phase at a certain location, the waves are said to interference constructively at that point. In this case, waves P and Q interfere constructively at all points along the x-axis (along the line of propagation). The superposition of these two waves produces a bigger resultant wave with the largest possible amplitude, $A+B$ — depicted by the dotted lines. As the resultant wave travels, the amplitude of oscillation of each particle attains the maximum value $A+B$. The coherency of the two waves ensures that they are lined up at all instances — otherwise, the amplitude of oscillation of each point will be a function of time and it would be meaningless to say if the amplitude is the largest or smallest.

Well, the larger amplitude above may seem obvious at first. After all, the addition of results of two things should intuitively be greater than the individual results. However, consider the second scenario where the phase difference between two waves is an odd multiple of π (i.e. $\Delta\phi = (2n+1)\pi$, $n \in \mathbb{Z}$) as depicted in Fig. 14.6.

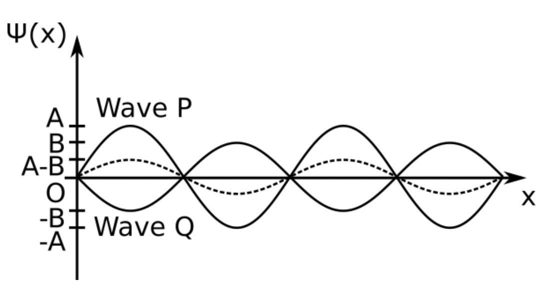


Figure 14.6: Destructive interference

Waves P and Q are said to be out of phase.[2] The superposition of these two waves produces a resultant wave with the smallest possible amplitude, $|A - B|$. Similar to the definition of constructive interference, waves are said to interfere destructively at a location if corresponding points at that location are out of phase. In this case, waves P and Q are said to undergo destructive interference at all points along the x-axis. In Fig. 14.6, the amplitude of the resultant wave (in dotted lines) is actually smaller than that of the individual waves! In fact, when the amplitudes of waves P and Q are equal, they will completely cancel out and the displacement will be zero in all space at all times. This is surprising in that if we had two speakers[3] that produce one-dimensional waves of a single frequency along the same line such that we currently hear sound of the greatest intensity between the speakers, positioning one speaker half a wavelength further from us would produce no sound at our location, though there are multiple sources of sound! In fact, destructive interference is widely applied in noise-cancellation systems which usually measure ambient sounds and produce sound waves that are out of phase with the ambient sounds to neutralize them.

We shall prove mathematically that the resultant amplitude of oscillation of a point due to two individual coherent waves is the largest and smallest when the phase difference between them is an even and odd multiple of π, respectively. It is assumed that the directions of displacement due to the two waves are identical. Let the two displacements at a particular point P be

$$\psi_1 = A\sin(kx + \omega t),$$
$$\psi_2 = B\sin(kx + \omega t + \Delta\phi).$$

Note that the waves only need to intersect at P — they do not necessarily overlap elsewhere. Again, the coherency of the two waves enables us to pick an origin in time for which the displacements are as above. The resultant

[2] The term "out of phase" shall be used to mean "perfectly out of phase."

[3] This set-up is actually not related to waves P and Q. It consists of two waves travelling in opposite directions and us measuring the intensity at one point between the two speakers. However, one interesting question to ponder is whether energy is conserved in this set-up when constructive or destructive interference occurs at our location. If the speakers individually produced waves of amplitude A at our location, the resultant wave will have amplitude $2A$ when constructive interference occurs — seemingly suggesting that the instantaneous energy at our location should be 4 times that which is individually produced by one speaker. However, we know from the conservation of energy that the instantaneous energy at our location should only be 2 times that of the individual one produced by each speaker. What is wrong here? The solution to this is presented later in this section.

displacement at P is

$$\begin{aligned}\psi_R &= A\sin(kx+\omega t) + B\sin(kx+\omega t+\Delta\phi)\\ &= A\sin(kx+\omega t) + B\sin(kx+\omega t)\cos\Delta\phi + B\sin\Delta\phi\cos(kx+\omega t)\\ &= (A+B\cos\Delta\phi)\sin(kx+\omega t) + B\sin\Delta\phi\cos(kx+\omega t).\end{aligned}$$

Applying the trigonometric R-formula,

$$\psi_R = \sqrt{A^2+B^2+2AB\cos\Delta\phi}\,\sin(kx+\omega t+\theta), \tag{14.5}$$

where $\theta = \tan^{-1}\frac{B\sin\Delta\phi}{A+B\cos\Delta\phi}$. The amplitude of ψ_R is the largest and smallest when $\cos\Delta\phi = 1$ and -1 respectively — corresponding to even and odd multiples of π respectively.

Finally, it is important to understand why constructive and destructive interferences produce the greatest and smallest intensities (e.g. loudest and softest sound or brightest and dimmest light) at a point, respectively. You may look at Fig. 14.5 and think that the location of the first non-zero x-intersect corresponds to zero intensity as the point there currently exhibits zero displacement. However, remember that the intensity of a wave at a location is proportional to the time-averaged squared displacement at that location. We have paused our movie in drawing Fig. 14.5 to observe the wave at a certain instance — resuming it would cause the resultant wave to continue traveling rightwards and the point at the first non-zero x-intersect to be displaced vertically. The final outcome is that the time-averaged displacement of that particular point is proportional to the squared amplitude of the wave at that point, $(A+B)^2$. Therefore, constructive interference between waves at a point corresponds to an intensity maximum — a similar logic holds for destructive inference.

We shall prove this claim mathematically. Consider two waves which individually produce displacements

$$\begin{aligned}\boldsymbol{D}_1 &= \boldsymbol{A}_1\cos(\omega t+\Delta\phi),\\ \boldsymbol{D}_2 &= \boldsymbol{A}_2\cos(\omega t),\end{aligned}$$

at a certain point P respectively. The phase of the first displacement leads that of the second by $\Delta\phi$ — this is not to be confused with wave Q leading wave P previously as that was a description of the entire wave. We write the displacements as vectors to accommodate the possibility of the planes of oscillations being different. The resultant vector displacement is

$$\boldsymbol{D}_r = \boldsymbol{D}_1 + \boldsymbol{D}_2.$$

The intensity I at point P is proportional to the time-averaged squared displacement at P over a single period. However, there is an important qualification to be made here. The previous statement is valid in the case of mechanical waves only if the two waves travel in the same direction and their displacements lie along the same line, but it is generally true for electromagnetic waves. This is because the kinetic energy of a medium generally does not directly scale with its displacement owing to the fact that we have to take into account the direction of motion. Yet, there is no kinetic energy term in the energy density of the electromagnetic wave — this is actually a hint to the resolution of the paradox presented in Footnote 3. If the angle between the individual vectorial displacements produced by the two sources is a constant θ (note that this is not necessarily the angle between the two rays joining the sources of the waves to P as electromagnetic waves are not longitudinal),

$$\begin{aligned}
I &\propto \langle D_r^2 \rangle \\
&= \langle (\boldsymbol{D}_1 + \boldsymbol{D}_2) \cdot (\boldsymbol{D}_1 + \boldsymbol{D}_2) \rangle \\
&= \langle D_1^2 \rangle + \langle D_2^2 \rangle + 2\langle \boldsymbol{D}_1 \cdot \boldsymbol{D}_2 \rangle \\
&= \frac{A_1^2}{2} + \frac{A_2^2}{2} + 2\langle A_1 A_2 \cos(\omega t + \Delta\phi) \cos(\omega t) \cos\theta \rangle \\
&= \frac{A_1^2}{2} + \frac{A_2^2}{2} + \langle A_1 A_2 \cos(2\omega t + \Delta\phi) \cos\theta \rangle + \langle A_1 A_2 \cos\Delta\phi \cos\theta \rangle \\
&= \frac{A_1^2}{2} + \frac{A_2^2}{2} + 0 + A_1 A_2 \cos\Delta\phi \cos\theta \\
&= \frac{A_1^2}{2} + \frac{A_2^2}{2} + A_1 A_2 \cos\Delta\phi \cos\theta.
\end{aligned}$$

When the directions of oscillation are aligned such that $\cos\theta = 1$,

$$I \propto \frac{1}{2}(A_1^2 + A_2^2 + 2A_1 A_2 \cos\Delta\phi),$$

where the term in brackets is the squared amplitude of the resultant wave that we have derived previously. Now, we can also express the resultant intensity in terms of the individual intensities of the waves. Let I_1 and I_2 be the individual intensities of the waves at P due to the two waves. Note that

$I_1 \propto \langle D_1^2 \rangle = \frac{A_1^2}{2}$ and $I_2 \propto \langle D_2^2 \rangle = \frac{A_2^2}{2}$. Then,

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \Delta\phi \cos \theta. \tag{14.6}$$

Evidently, the maximum and minimum intensities occur at P for acute θ when the conditions for constructive and destructive interferences are satisfied respectively at P (so that $\cos \Delta\phi = 1$ and -1).[4] In the case of transverse waves, such as light waves, the directions of oscillations are usually aligned — causing $\cos \theta = 1$ and the above equation to become

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \Delta\phi.$$

Let us digress for a bit to resolve an apparent paradox that befuddles many. In the case of constructive interference of the previous waves P and Q when they have the same amplitude and no phase difference, the resultant wave is a version of wave P or Q, scaled by a factor of 2. Therefore, one may think that we have obtained a contradiction as the instantaneous energy of each point on the resultant wave has 4 times its corresponding energy on each component wave — thus "violating the conservation of energy" as we expect the instantaneous energy to be the sum of that produced by the two component waves. Similarly, when waves P and Q interfere destructively to annihilate each other completely, the instantaneous energy of each point seemingly vanishes, though each point has a certain instantaneous energy on the component waves! Well, the resolution to this paradox is that the real physical wave is given by the superposition of waves P and Q. Therefore, it is correct to say that the instantaneous energies are 4 times and 0 times that produced by each component wave above. However, there is no meaning in ascribing energy to the component waves as the current physical wave is neither of the component waves. It is correct to say that each component wave results in a certain instantaneous energy when they are present alone but it is incorrect to say that the instantaneous energy of the resultant wave produced by their superposition is the addition of their individual instantaneous energies, as energy is not a linear function. Perhaps, the best illustration of this fact is that the driver of the resultant wave delivers power according to the resultant wave at its location and not the sum of the powers of the component waves.

But wait! What if the component waves have yet to reach the drivers such that the drivers do not yet know of their existence? For such a case to exist, the wave must travel in different directions. An example would

[4] Note that if θ is obtuse, the conditions are flipped as $\cos \theta$ is now negative.

be the two identical string pulses produced by two far away sources that travel in opposite directions in Fig. 14.1. Persons A and B each produce a wave pulse carrying a certain amount of energy, without regard for the wave produced by the other. Is energy still conserved when the two wave pulses overlap in this case? Well, the answer had better be yes and we will indeed reach this conclusion if we consider the different components of energy meticulously. For example, you might think that the total potential energy of the resultant wave when the two pulses completely overlap (see Fig. 14.2) is 4 times that of each individual pulse (as the amplitude doubles such that the gradient at each point and the amount that each section of string is stretched by, follows suit) and thus claim that the conservation of energy is violated. However, under closer scrutiny, you would notice that the kinetic energy of the resultant pulse at this juncture is zero but the component waves originally possessed a certain amount of kinetic energy each. In fact, we know from the previous chapter that for a traveling wave, the instantaneous kinetic and potential energies of a section of string are identical! Therefore, the kinetic energies of the individual pulses (which constitute twice the potential energy of an individual pulse) have been converted to potential energy[5] to produce a resultant pulse with 4 times the potential energy of an individual pulse at the juncture where the two pulses completely coincide.

To show the conservation of energy in the general case (when the pulses do not completely overlap and are non-identical), we can exploit the result of Problem 4 in Chapter 14 which states that the power delivered across a point on a mechanical wave, produced by the superposition of two waves traveling in opposite directions, is simply the sum of the individual powers delivered by each component wave. That is, in the overlapping region between the two waves,

$$P(x,t) = P_1(x,t) + P_2(x,t),$$

[5]This is in fact the solution to the question posed in Footnote 3 as the instantaneous kinetic energies carried by the two component waves produced by the speakers have been supplemented to the potential energy of the air section at our location, when constructive interference occurs. That is, the potential energy at our location is 4 times the individual one that would have been caused by each wave alone but there is still a local conservation of energy as the kinetic energy of the resultant wave at our location is zero. In conclusion, both the total energy density and amplitude at our location doubles as compared to those produced by a single speaker. In retrospect, the fallacy in Footnote 3 stems from the claim that the intensity at a point is proportional to its squared amplitude for a superposition of mechanical waves traveling in different directions, when it is only valid for a single traveling wave.

where $P(x,t)$, $P_1(x,t)$ and $P_2(x,t)$ are the instantaneous powers delivered from the left to the right of a point at equilibrium x-coordinate x by the resultant wave and the first and second component waves respectively. The rate of increase of the linear energy density $\varepsilon(x,t)$ (which includes both kinetic and potential energies) carried by a section of string between equilibrium coordinates x and $x+dx$ is thus

$$\frac{\partial \varepsilon(x,t)}{\partial t}dx = P(x,t) - P(x+dx,t),$$
$$\frac{\partial \varepsilon}{\partial t} = -\frac{\partial P}{\partial x} = -\frac{\partial P_1}{\partial x} - \frac{\partial P_2}{\partial x},$$

where $-\frac{\partial P_1}{\partial x}$ and $-\frac{\partial P_2}{\partial x}$ are the individual rates of change that would have been engendered by the first and second component waves. The above equation hence encapsulates the conservation of energy; it states that the change in energy of a section is equal to the sum of the individual powers delivered to it by the two component waves. In fact, it is a stronger statement than the conservation of energy as it implies that conservation of energy is local. That is, not only is energy conserved as a whole, the net energy entering into one section from its neighbours must become its increase in mechanical energy such that there is a continuity of energy. Energy cannot suddenly disappear from one end of a wave and appear at another (which results in energy conservation that is not local).

Returning to our main topic, waves may generally travel in different directions and cross at a particular point.[6] In light of the fact that our previous equations for intensity are only valid for electromagnetic waves when the waves travel in different directions, all sources in the rest of this section will only refer to light sources. As seen from our previous analysis, the phase differences between different waves at a particular point are factors in determining the intensity at that point. Only the interference between two wave sources will be considered for now. As a recap, spatial positions at which the phase difference between corresponding points on the waves is an even multiple of π are known as regions of constructive interference while those at which the phase difference is an odd multiple of π are defined to be regions of destructive interference. Since the two sources are assumed to be coherent (i.e. their phase differences at these locations are invariant with respect to time), regions of constructive and destructive interferences remain so at all times.

[6] Only waves whose paths intersect are of interest to us. If their paths do not intersect, such as in the case of skew lines, no interference occurs — causing them to be uninteresting.

The phase difference between two waves may arise from the different paths they travel as the phase of a point on a wave varies with its position at a fixed time. We shall now derive the conditions for constructive and destructive interferences to occur at a particular point in space P due to coherent light sources.

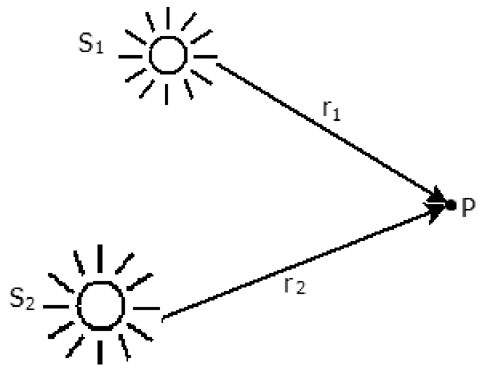


Figure 14.7: Interference between two sources at P

Suppose that the waves leaving the two sources are initially in phase and that the relevant θ (angle between the directions of oscillation) is acute. At point P, the waves from sources S_1 and S_2 would have traveled path lengths of r_1 and r_2 respectively. There is a path difference of $\delta = r_1 - r_2$. Since the phase angle of a point on a wave at a particular instant increases by 2π for every additional distance λ along the wave by Eq. (14.4), the condition for constructive interference to occur at point P is

$$\delta = r_1 - r_2 = n\lambda. \qquad (n \in \mathbb{Z})$$

Then, the condition for destructive interference to occur at point P is

$$\delta = r_1 - r_2 = \left(n - \frac{1}{2}\right)\lambda. \qquad (n \in \mathbb{Z})$$

If the two sources are not perfectly in phase, the above conditions have to be modified. If we let source S_1 lead source S_2 by a phase of ϕ_o, the waves at point P due to S_1 have essentially traveled an additional path length of $\frac{\phi_o}{2\pi}\lambda$. The set-up is identical to a hypothetical experiment in which a source S_1' is placed $\frac{\phi_o}{2\pi}\lambda$ behind[7] S_1. Thus, the conditions for constructive and destructive

[7] We take the direction of S_1 to P to be the forward direction.

interferences become

$$r_1 + \frac{\phi_o}{2\pi}\lambda - r_2 = n\lambda, \qquad (n \in \mathbb{Z})$$

$$r_1 + \frac{\phi_o}{2\pi}\lambda - r_2 = \left(n - \frac{1}{2}\right)\lambda, \qquad (n \in \mathbb{Z})$$

respectively. To determine the resultant intensity at P in terms of the individual intensities, observe that the phase difference between the waves produced by S_1 and S_2 is $\Delta\phi = k(r_1 - r_2) + \phi_o$. Substituting this expression for $\Delta\phi$ into Eq. (14.6) would yield the desired result.

14.1.2 *Young's Double Slit Experiment*

Thomas Young demonstrated the wave nature of light with his famous double slit experiment. A monochromatic light source S is placed in front of a board with a small slit which is presumed to be long (into the page). Source S illuminates the board with plane waves of wavelength λ, which is physically attainable by placing the source very far away. Huygen's principle asserts that every unobstructed point on a wave front is a point source of spherical waves which interfere with each other. By Huygen's principle, we can consider each point on the slit to be a spherical source (due to the transmitted wave fronts). Since the slit S_0 is long, the spherical waves produced by each point on the slit interfere to yield an essentially cylindrical wave. Hence, we can ignore the direction perpendicular to the plane of the page in the following analysis.

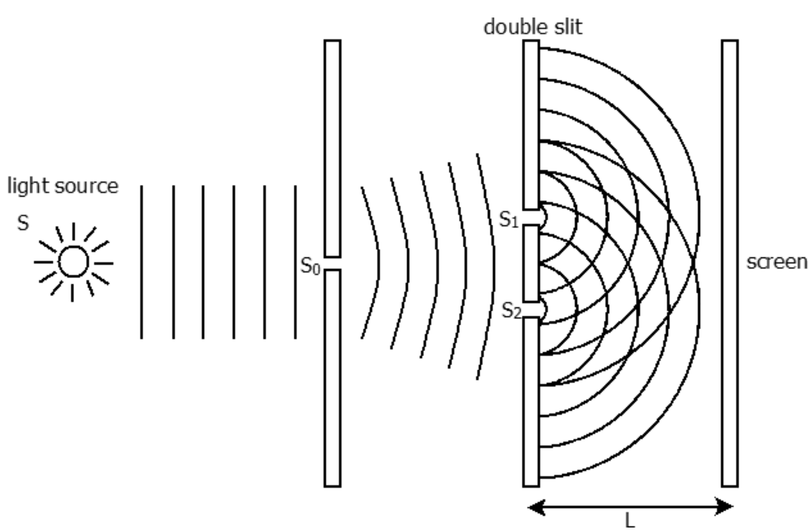


Figure 14.8: Double slit experiment

Next, another board, with two long slits of infinitesimal thickness that are parallel to the first slit, is positioned symmetrically about the vertical position of S_0 — they are separated by a distance d. The two infinitesimal slits S_1 and S_2 then develop into two sources of cylindrical waves that are coherent. Furthermore, these two sources are in phase, assuming that S_1 and S_2 are equidistant from S_0 and that the plane waves emitted by S are parallel to the first board.

The light waves produced by the two cylindrical sources S_1 and S_2 interfere and produce an interference pattern as they impinge on a distant screen. How should this pattern look like? There should be alternating bright and dark regions due to progressive changes in the phase difference between the rays emanating from the slits that impinge on the screen. We can first identify the regions of constructive and destructive interferences which correspond to regions of maximum and minimum amplitudes and thus, intensities.

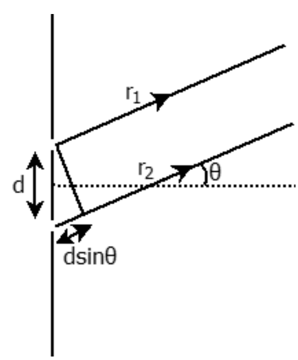


Figure 14.9: Far field approximation

Before we adopt the most rigorous method of summing the individual contribution of each source, observe that in the far field scenario in which the screen is distant from the two slits (explicitly, $d \ll L$ where L is the horizontal distance between the two slits and the screen), the paths taken by two waves, which originate from S_1 and S_2 and coincide at the same position on the screen, are essentially parallel. Furthermore, the ratio of the distances traveled by rays from the sources to the same point on the screen is essentially 1 ($\frac{r_1}{r_2} \approx 1$). However, the additive difference $r_2 - r_1$ is non-negligible as we will compare it to λ, which is small as well, when we are determining the conditions for constructive and destructive interferences. If θ is the angle that these parallel rays make with the horizontal and d is the vertical distance between the two slits, the path difference between the two waves is (see Fig. 14.9)

$$\delta = r_2 - r_1 = d \sin\theta.$$

Therefore, the values of θ at which constructive and destructive interferences occur are

$$d \sin\theta = n\lambda, \qquad (n \in \mathbb{Z})$$

$$d \sin\theta = \left(n - \frac{1}{2}\right)\lambda, \qquad (n \in \mathbb{Z})$$

respectively. At points of constructive and destructive interferences, the resultant amplitude will be a local maximum and minimum correspondingly. Since the intensity[8] at a point is proportional to the squared amplitude at that point, regions of constructive and destructive interferences correspond to bright and dark fringes respectively.

At points where neither constructive nor destructive interference occurs, their intensities attain intermediate values. Therefore, the intensities of various points on the wall at different vertical positions can be plotted as in Fig. 14.10.

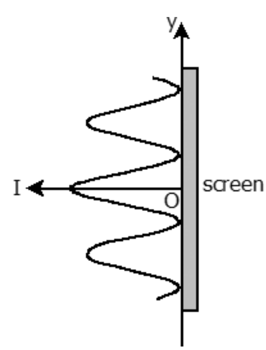


Figure 14.10: Intensity pattern

The positions of bright and dark fringes correspond to the local intensity maxima and minima respectively. The intensities of the bright fringes are lower the further they are from the center. This is due to the decreasing amplitudes[9] of the waves as they traverse longer distances with increasing θ. However, note that the waves almost annihilate completely at the intensity

[8]Take note that the brightness of a point is quantified by the intensity, rather than the amplitude, at that point.

[9]The amplitude of a cylindrical wave decreases as $\frac{1}{\sqrt{r}}$, where r is the distance from the source.

minima as the distances traveled by the waves are essentially equal in the multiplicative sense ($\frac{r_2}{r_1} \approx 1$), for a given θ in the far field approximation.

The quantity n in the equations, describing the conditions for constructive and destructive interferences above, indicates the $|n|$th order bright and dark fringes, or $|n|$th order maximum and minimum, respectively. For example, substituting $n = 0$ in the equation for constructive interference governs the zeroth-order maximum, or bright fringe which is at the center. The two adjacent bright fringes are then first-order bright fringes.

Next, we can determine the approximate vertical positions of the bright and dark fringes when θ is small by using the small angle approximation

$$\sin\theta \approx \tan\theta \approx \frac{y}{L},$$

where y is the vertical coordinate with respect to the origin defined at the vertical center of the two slits and L is the distance between the screen and the slits. Substituting this expression for $\sin\theta$ in the equations for constructive and destructive interferences yields

$$y_{bright} = n\frac{\lambda L}{d}, \qquad (n \in \mathbb{Z})$$

$$y_{dark} = \left(n - \frac{1}{2}\right)\frac{\lambda L}{d}. \qquad (n \in \mathbb{Z})$$

Most notably, the above equations imply that the separations between successive bright fringes and dark fringes are both

$$\Delta y = \frac{\lambda L}{d}.$$

Problem: Why must the source be monochromatic for bright and dark fringes to be clearly observed?

If the light consists of various wavelengths, the interference pattern of each wavelength of light would involve bright and dark fringes at different locations on the screen. The mashing of these interference patterns via the principle of superposition would obscure the interference pattern.

Problem: If the monochromatic light waves emitted by the slit on top initially lead those emitted by the one on the bottom by a phase of $\frac{\pi}{2}$, and given that a second-order maximum occurs at a positive vertical coordinate y_2 from the center of the screen, what is the wavelength of the light? The origin is defined to be at the same position as before, and $y_2 \ll L$ where L is the horizontal distance between the screen and the two slits.

The $\frac{\pi}{2}$-radian phase lead is essentially equal to an additional distance $\frac{\lambda}{4}$ traveled by the waves emanating from the top slit. Therefore, for constructive interference, the path difference must satisfy

$$d\sin\theta - \frac{\lambda}{4} = n\lambda$$
$$\implies d\sin\theta = \left(n + \frac{1}{4}\right)\lambda.$$

By the small angle approximation,

$$\sin\theta \approx \tan\theta \approx \frac{y}{L}.$$

Substituting $y = y_2$ when $n = 2$ and solving,

$$\lambda = \frac{4y_2 d}{9L}.$$

Intensity

Label each point on the distant screen with a coordinate θ which denotes the angle subtended by a line, joining the center of the slits to that point on the screen, and the horizontal. The intensity profile on the screen $I_{tot}(\theta)$ can be determined by applying Eq. (14.6) (with $\theta = 0$ in the equation[10] as the waves from the slits originated from the same source).

$$I_{tot}(\theta) = I_1(r_1) + I_2(r_2) + 2\sqrt{I_1(r_1)I_2(r_2)}\cos(kd\sin\theta),$$

where the phase difference $\Delta\phi = \frac{2\pi}{\lambda}d\sin\theta = kd\sin\theta$ in this case. I_1 and I_2 are the intensities that would have been produced by the waves emanating from the slits and landing on the point on the screen corresponding to θ, individually. They are functions of the path lengths traversed by their respective rays, r_1 and r_2. Concretely,

$$I_1(r_1) \propto \frac{1}{r_1} \quad \text{and} \quad I_2(r_2) \propto \frac{1}{r_2}$$

for cylindrical waves. As $\frac{r_1}{r_2} \approx 1$ and the individual intensities at the point corresponding to $\theta = 0$ are identical, the individual intensities are essentially equal for a given θ. Then, we can represent them in terms of a common

[10]Do not confuse this with the θ coordinate in this section.

intensity that is a function of θ, $I(\theta)$.

$$I_1(r_1) = I_2(r_2) = I(\theta).$$

Observe that with this new definition,

$$I(\theta) = I(0)\cos\theta,$$

where $I(0)$ is the individual intensity at the center of the screen as the ratio of the path length covered by a ray to $\theta = 0$ and that to θ is $\cos\theta$. Thus,

$$I_{tot}(\theta) = 2I(0)\cos\theta(1 + \cos kd\sin\theta) = 2I(0)\cos\theta\cos^2\left(\frac{kd\sin\theta}{2}\right).$$

Usually, we compare $I_{tot}(\theta)$ to the net intensity at the center of the screen $I_{tot}(0)$ to visualize the shape of the intensity profile.

$$\frac{I_{tot}(\theta)}{I_{tot}(0)} = \cos\theta\cos^2\left(\frac{kd\sin\theta}{2}\right). \tag{14.7}$$

It is evident from this expression that the intensity maxima correspond to points where

$$\frac{kd\sin\theta}{2} = n\pi \qquad (n \in \mathbb{Z})$$
$$\implies d\sin\theta = n\lambda.$$

The intensity minima occur when

$$\frac{kd\sin\theta}{2} = \left(n - \frac{1}{2}\right)\pi \qquad (n \in \mathbb{Z})$$
$$\implies d\sin\theta = \left(n - \frac{1}{2}\right)\lambda,$$

which confirms our conditions obtained from the previous heuristic method.

14.1.3 *Diffraction Gratings*

A diffraction grating consists of many fine, parallel slits that are separated by a standardized distance, commonly inscribed on a sheet of glass or metal. A typical diffraction grating has several hundreds to thousands of slits per mm, resulting in a slit distance d which is much smaller than that in double slit

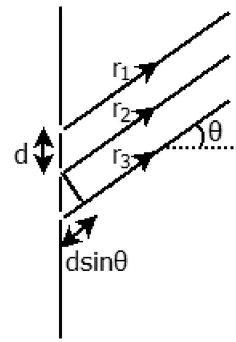


Figure 14.11: Diffraction grating

experiments. Usually, N, the number of slits per mm of a diffraction grating, is given. Then, the slit separation can be calculated as

$$d = \frac{1}{N}.$$

Again, we consider the ideal scenario where each slit is infinitely long but infinitesimally thin. As always, the slits act as cylindrical sources that are in phase when the grating is illuminated with plane waves whose wave fronts are parallel to the grating.

The monochromatic waves from each slit overlap to produce an interference pattern on a screen that is a distance L away. In the far field approximation, the rays — each originating from a single slit and coinciding at the same point P on the screen — essentially propagate along parallel lines. Then, the path difference between adjacent rays is still $\delta = d \sin\theta$.

Evidently, all of these waves interfere constructively when

$$\delta = d \sin\theta = n\lambda \qquad (n \in \mathbb{Z})$$

as they will all be in phase at that particular point P on the screen — bright fringes are produced on points on the screen that correspond to such θ. However, these are not the only local intensity maxima on the screen — there are also low-intensity maxima (this is to be expected as we only considered the ideal case where all rays interfere constructively and left out other combinations). Next, the previous condition for destructive interference cannot be used to determine the locations of dark fringes here for the following reason. Suppose that the top ray annihilated the middle ray in the above diffraction grating consisting of three slits — the bottom ray will still survive. Ultimately, the cancellation of pairs of waves does not guarantee the minimization of the effect of the group.

Intensity

In this section, we will determine the intensity profile $I_{tot}(\theta)$ due to M slits. Before we embark on this goal, let us introduce the following idea which is rather general. Recall from the previous chapter that a traveling wave of the form $\psi = A\cos(kr - \omega t + \phi)$ can be represented in complex form as

$$\tilde{\psi} = Ae^{i(kr-\omega t+\phi)}.$$

Then, linear operations can be performed on this complex representation — the real, physical result can then be retrieved by taking the real component of the final complex result. Since the planes of oscillations of the electric fields due to the slits are aligned, the net complex electric field at a point on the screen, $\tilde{E}_{tot}(\theta)$, is given by the sum of the individual complex electric fields, i.e.

$$\tilde{E}_{tot}(\theta) = \sum_{j=1}^{M} A_j(r_j)e^{i(kr_j-\omega t+\phi)},$$

where $A_j(r_j)$ is the real amplitude of the wave produced by the jth slit on the point on the screen corresponding to angle θ and r_j is the path length traversed by that wave front from the jth slit in doing so. ϕ is a constant phase offset which is uniform across all slits as we assume that the waves emanating from the slits are initially in phase. Finally, the real net electric field can be obtained from the real component of $\tilde{E}_{tot}(\theta)$. Since the ratios of pairs of r_j's are all essentially 1 for a given θ and because $A_j \propto \frac{1}{\sqrt{r_j}}$ for a cylindrical wave,

$$A_1(r_1) \approx A_2(r_2) \approx \ldots \approx A_M(r_M),$$

and we can instead define $A(\theta)$ as the common amplitude that is solely dependent on θ. Then,

$$\tilde{E}_{tot}(\theta) = A(\theta)e^{-i(\omega t-\phi)}\sum_{j=1}^{M} e^{ikr_j}.$$

The real net electric field is then

$$E_{tot}(\theta) = \mathrm{Re}\left[\tilde{E}_{tot}(\theta)\right] = A(\theta)\cos(\omega t - \phi)\mathrm{Re}\left[\sum_{j=1}^{M} e^{ikr_j}\right]$$
$$+ A(\theta)\sin(\omega t - \phi)\mathrm{Im}\left[\sum_{j=1}^{M} e^{ikr_j}\right].$$

For the sake of convenience, denote x as $\mathrm{Re}\left[\sum_{j=1}^{M} e^{ikr_j}\right]$ and y as $\mathrm{Im}\left[\sum_{j=1}^{M} e^{ikr_j}\right]$. The net intensity $I(\theta)$ is proportional to the time average of the squared electric field.

$$\begin{aligned} I_{tot}(\theta) &\propto \langle E_{tot}^2(\theta)\rangle \\ &= \langle A^2(\theta)\cos^2(\omega t-\phi)x^2\rangle + \langle A^2(\theta)\sin^2(\omega t-\phi)y^2\rangle \\ &\quad + \langle 2A^2(\theta)\sin(\omega t-\phi)\cos(\omega t-\phi)xy\rangle \\ &= \frac{1}{2}A^2(\theta)(x^2+y^2)+0 \\ &\propto A^2(\theta)\left|\sum_{j=1}^{M} e^{ikr_j}\right|^2, \end{aligned}$$

as x^2+y^2 simply yields the squared magnitude of $\sum_{j=1}^{M} e^{ikr_j}$. At this point, we can substitute our expression for $A(\theta)$ in terms of $A(0)$. For cylindrical waves,

$$A(\theta) = A(0)\sqrt{\cos\theta},$$

with the repercussion that

$$I_{tot}(\theta) \propto \cos\theta\left|\sum_{j=1}^{M} e^{ikr_j}\right|^2.$$

Therefore, the remaining task is to determine $|\sum_{j=1}^{M} e^{ikr_j}|^2$. To tackle this, observe that $r_{j+1}-r_j = d\sin\theta$ for any $1 \le j \le M-1$. Consequently,

$$\begin{aligned} \sum_{j=1}^{M} e^{ikr_j} &= e^{ikr_1}\sum_{j=1}^{M} e^{i(j-1)kd\sin\theta} \\ &= e^{ikr_1}\cdot\frac{1-e^{iMkd\sin\theta}}{1-e^{ikd\sin\theta}} \\ &= e^{ikr_1}\cdot\frac{e^{i\frac{Mkd\sin\theta}{2}}}{e^{i\frac{kd\sin\theta}{2}}}\cdot\frac{e^{-i\frac{Mkd\sin\theta}{2}}-e^{i\frac{Mkd\sin\theta}{2}}}{e^{-i\frac{kd\sin\theta}{2}}-e^{i\frac{kd\sin\theta}{2}}} \\ &= e^{i\left(kr_1+\frac{(M-1)kd\sin\theta}{2}\right)}\frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{\sin\left(\frac{kd\sin\theta}{2}\right)}, \end{aligned}$$

where we have applied the geometric series formula in writing the second equality. It can be seen that the squared magnitude of the above is

$$\left|\sum_{j=1}^{M} e^{ikr_j}\right|^2 = \frac{\sin^2\left(\frac{Mkd\sin\theta}{2}\right)}{\sin^2\left(\frac{kd\sin\theta}{2}\right)}$$

$$\implies I_{tot}(\theta) \propto \cos\theta \frac{\sin^2\left(\frac{Mkd\sin\theta}{2}\right)}{\sin^2\left(\frac{kd\sin\theta}{2}\right)}.$$

To determine the ratio between $I_{tot}(\theta)$ and $I_{tot}(0)$, we have to evaluate

$$\frac{I_{tot}(\theta)}{I_{tot}(0)} = \cos\theta \frac{\sin^2\left(\frac{Mkd\sin\theta}{2}\right)}{\sin^2\left(\frac{kd\sin\theta}{2}\right)} \cdot \lim_{\theta\to 0} \frac{\sin^2\left(\frac{kd\sin\theta}{2}\right)}{\sin^2\left(\frac{Mkd\sin\theta}{2}\right)}.$$

Since $\sin x \approx x$ for small x, the last term yields $\frac{1}{M^2}$.

$$\frac{I_{tot}(\theta)}{I_{tot}(0)} = \cos\theta \left(\frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{M\sin\left(\frac{kd\sin\theta}{2}\right)}\right)^2, \tag{14.8}$$

which is a rather neat result. Since the minima occur when $I_{tot}(\theta) = 0$, you may think that the relevant condition is

$$\frac{Mkd\sin\theta}{2} = n\pi$$

$$\implies d\sin\theta = \frac{n}{M}\lambda.$$

However, this is not entirely correct for the cases when $d\sin\theta = n'\lambda$ for some integer n' as the denominator also tends to zero. Such scenarios in fact result in the high-intensity maxima as argued heuristically before. Otherwise, one can also use the Taylor expansion $\sin(n\pi + x) \approx x$ to show that

$$\lim_{\frac{kd\sin\theta}{2}\to n'\pi} \frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{\sin\left(\frac{kd\sin\theta}{2}\right)} = \frac{\frac{Mkd\sin\theta}{2} - Mn'\pi}{\frac{kd\sin\theta}{2} - n'\pi} = M.$$

This is in fact the global maximum of the above expression. Correspondingly, $\frac{I_{tot}(\theta)}{I_{tot}(0)} = \cos\theta$ in such cases. The $\cos\theta$ term only acts as a slowly-varying envelope for the wildly oscillating[11] $\left(\frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{M\sin\left(\frac{kd\sin\theta}{2}\right)}\right)^2$ term so it does not affect the positions of the extrema[12] significantly. Therefore, cases where

[11] d is usually a few orders of magnitude above λ such that kd is significant.

[12] The $\cos\theta$ term slightly affects the positions of the maxima but not the minima (where the net intensity is zero).

$d \sin\theta = n'\lambda$ correspond to high-intensity peaks while the intensity minima occur when

$$d \sin\theta = \frac{n}{M}\lambda \quad | \quad n \nmid M.$$

Evidently, there are $(M-1)$ minima between successive high-intensity maxima. Finally, the $d\sin\theta$ values of the low-intensity maxima are in fact situated approximately in between those of two adjacent intensity minima for large M (i.e. $d\sin\theta = \frac{2n+1}{2M}$) — aggregating to yield $(M-2)$ low-intensity peaks between two neighboring high-intensity peaks.

If we plot $I_{tot}(\theta)$ for $M = 3$ slits, we obtain the following graph.

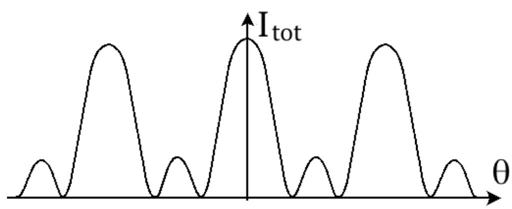


Figure 14.12: Intensity graph

It can be seen that the exact locations of the intensity minima are not particularly enlightening for the following reason. Due to the relatively large distances between high-intensity peaks and the preponderance of low-intensity maxima (small bumps) between high-intensity peaks, the locations of dark fringes are usually neglected as the width of the low-intensity region, which includes both low-intensity maxima and intensity minima, is large. The low-intensity maxima and the intensity minima are often indistinguishable when M is large.

Lastly, there are certain differences between the interference pattern produced by a diffraction grating and a board with two slits.

- As d is often smaller for a diffraction grating, the angle θ for bright fringes is usually much larger. Hence, the small angle approximation $\sin\theta \approx \frac{y}{L}$ does not hold. It is not meaningful to determine the fringe separation between consecutive high-intensity peaks as it is not a constant. In fact, it increases as the order increases.
- As the number of slits M increases, the high-intensity maxima (bright fringes) become narrower and appear sharper on the screen. Quantitatively, for small θ, $I_{tot}(\theta)$ is periodic for every 2π-increment of $kd\sin\theta$. Therefore, the width of the central bright fringe is a good representation

of the neighboring ones. Its angular width is twice the angle θ corresponding to the first-order minima and thus $\frac{2\lambda}{Md}$. Evidently, this angular width decreases with increasing M — causing the fringe to be concentrated. Thus, the contrast of the interference pattern becomes more stark.

- Interestingly, for small values of $kd\sin\theta$, the ratio $\frac{I_{tot}(\theta)}{I_{tot}(0)}$ does not vary with M — implying that the relative brightnesses of the fringes do not change even though the absolute intensity of the entire profile increases with M (as more light passes through a larger number of slits).

Resolving Power

Since a diffraction grating is often applied to separate a bundle of light with different wavelengths into a spectrum, a rather practical question to ask is: what is the minimum value of $\Delta\lambda > 0$ such that we are able to distinguish between the nth order maxima of the waves with wavelengths λ and $\lambda + \Delta\lambda$? A widely accepted standard for the resolvability of two objects is that the peak of the interference pattern of one image lies beyond the adjacent minimum of the other image — this is known as the Rayleigh criterion. Applied to this situation, the boundary case occurs[13] when the nth order high-intensity maximum of $\lambda + \Delta\lambda$ coincides with the $(nM+1)$th order minimum of λ (the minimum nearest to the nth order bright fringe of the other wavelength). The respective conditions for these are

$$\begin{aligned} d\sin\theta &= n(\lambda + \Delta\lambda) \\ d\sin\theta &= \frac{nM+1}{M}\lambda \\ \implies 1 + \frac{\Delta\lambda}{\lambda} &= \frac{nM+1}{nM}. \end{aligned}$$

The (spectral) resolving power R of an optical apparatus is defined as

$$R = \frac{\lambda}{\Delta\lambda} = nM.$$

A larger resolving power for fixed λ indicates a greater ability in resolving features of different wavelengths (as $\Delta\lambda$ decreases). The resolving power of a diffraction grating hence increases with the number of slits M and the order of the observed high-intensity maximum.

[13] One may wonder why we do not favor the boundary case where the nth order high-intensity maximum of λ overlaps with the $(nM-1)$th order minimum of $\lambda + \Delta\lambda$. Proceeding with the same process in this scenario would show that the condition in our current case is slightly stricter (though not significantly when M is large).

14.1.4 *Fraunhofer Diffraction*

Diffraction refers to the wave pattern formed when a wave scatters around a small aperture, whose width is comparable to the wavelength of the wave. As a wave is incident on a hole on a wall, some portions are absorbed or reflected by the wall while the others pass through the hole safely. From the perspective of Huygen's principle, diffraction is quintessentially the interference of the spherical waves produced by points on the wave fronts that are transmitted.

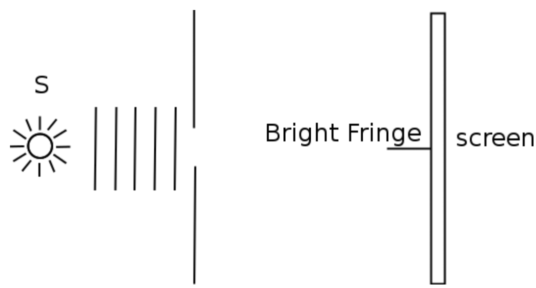


Figure 14.13: Incoming plane waves

In light of the above discussion, a single slit of non-negligible width can also exhibit an interference pattern. Consider a set of plane waves that pass through a long aperture of non-negligible width w. Every thin line on the slit (extending into the plane of the paper) acts as a cylindrical source which interferes with one another — thus producing bright and dark regions on a screen that is far away.

Observe that two point sources on the slit that are equidistant from the midpoint are equidistant from the center of the screen — implying that they interfere constructively there. This occurs for all pairs of points on the slit that are equidistant from the midpoint. Though the waves due to different pairs have slightly different phases at the center of the screen, the overall effect still produces a bright fringe. Unfortunately, the general locations of intensity maxima other than the center are difficult to determine without rigorously summing the contributions from each point source. However, it turns out that the magnitude of the other intensity maxima are negligible as we shall see later — implying that we are not at a real disadvantage here.

The dark fringes can thankfully be generally determined via the following argument. Divide the slit into two equal halves and consider two points A and B on the top and bottom halves. The distance between A and the

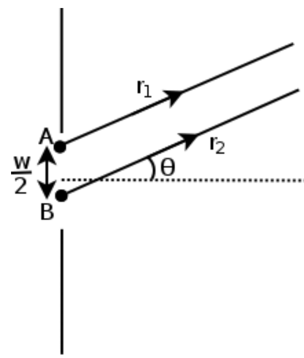


Figure 14.14: Two equal segments

top of the slit is equal to that between B and the midpoint such that the distance between A and B is $\frac{w}{2}$ (Fig. 14.14). Again, we are considering the far field approximation where the rays from points on the slit to the screen are parallel — this requires the distance between the slit and the screen to be much larger than w. The diffraction pattern under such conditions is known as Fraunhofer diffraction.

Consider a point on the screen at which parallel rays from A and B subtend an angle θ with respect to the horizontal. We understand from the previous analysis that the waves from A and B interfere destructively when

$$\frac{w}{2}\sin\theta = \pm\frac{\lambda}{2}.$$

We do not include values such as $\pm\frac{3\lambda}{2}$, $\pm\frac{5\lambda}{2}, \ldots$ on the right-hand side as they will be encapsulated later. The above destructive interference occurs for all pairs of corresponding points $\frac{w}{2}$ apart — as θ is identical with respect to all points on the slit in the far field approximation. Therefore, the point on the screen corresponding to such an angle θ is a dark fringe.

A similar analysis can be performed for any division of the slit into an even number of segments. For example, we can divide the slit into four and consider four corresponding points $\frac{w}{4}$ apart (Fig. 14.15).

Again, destructive interference occurs between the waves produced by the two adjacent pairs A, B and C, D at angle θ when

$$\frac{w}{4}\sin\theta = \pm\frac{\lambda}{2},$$

and holds for all such quadruples. In the general case where we split the slit into $2n > 0$ segments, the destructive interference condition is

$$\frac{w}{2n}\sin\theta = \pm\frac{\lambda}{2}.$$

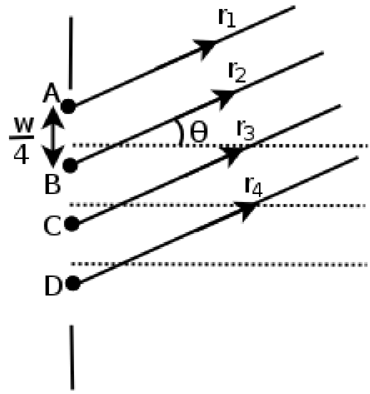


Figure 14.15: Four equal segments

The overall condition for destructive interference within a single slit is thus

$$w\sin\theta = n\lambda, \tag{14.9}$$

where n is a non-zero integer.

Intensity

We can extend the method introduced in the previous section for a diffraction grating to calculate the intensity pattern $I_{tot}(\theta)$ produced by a single slit. Instead of evaluating $\left|\sum_{j=1}^{M} e^{ikr_j}\right|^2$, we now have to compute the continuous integral $\left|\int e^{ikr(x)}dx\right|^2$ where $r(x)$ denotes the path length traversed by a wave emanating from coordinate x on the slit (the x-axis is defined to be positive upwards along the slit) and ending at the point on the screen corresponding to angle θ, as each infinitesimal section of the slit acts as a source of cylindrical waves. Defining our origin at the center of the slit (any origin would work) and observing that the path difference between a wave from coordinate x and one from the origin is $x\sin\theta$, we have

$$\begin{aligned}\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ikr(x)}dx &= e^{ikr(0)}\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ikx\sin\theta}dx \\ &= e^{ikr(0)}\left[\frac{e^{ikx\sin\theta}}{ik\sin\theta}\right]_{-\frac{w}{2}}^{\frac{w}{2}}\end{aligned}$$

$$= e^{ikr(0)} \frac{e^{i\frac{kw\sin\theta}{2}} - e^{-i\frac{kw\sin\theta}{2}}}{ik\sin\theta}$$
$$= e^{ikr(0)} \frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}.$$

The squared amplitude of this is evidently $\left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}\right)^2$. Piecing this component together with the $\cos\theta$ factor associated with the decreasing amplitude of $A(\theta)$ for cylindrical waves,

$$I_{tot}(\theta) \propto \cos\theta \left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}\right)^2,$$

$$\frac{I_{tot}(\theta)}{I_{tot}(0)} = \cos\theta \left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}\right)^2 \div \lim_{\theta\to 0} \left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}\right)^2.$$

We can use[14] the small angle approximation $\sin x \approx x$ to prove that the last term is simply w^2. Hence,

$$\frac{I_{tot}(\theta)}{I_{tot}(0)} = \cos\theta \left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{kw\sin\theta}{2}}\right)^2. \tag{14.10}$$

For small angles of θ (this assumption will soon be justified),

$$\frac{I_{tot}(\theta)}{I_{tot}(0)} = \left(\frac{\sin\alpha}{\alpha}\right)^2 = \text{sinc}^2\,\alpha, \tag{14.11}$$

where a new variable $\alpha = \frac{kw\sin\theta}{2}$ has been introduced and the function $\text{sinc}\,\alpha$ represents $\frac{\sin\alpha}{\alpha}$. The function $\text{sinc}^2\,\alpha$ is plotted in Fig. 14.16.

It can be seen that the intensity plunges rapidly as the $\frac{1}{\alpha}$ term decays the amplitude of $\sin\alpha$. Most of the total energy is concentrated within the central bright fringe — justifying our analysis of small θ as the $\cos\theta$ term would only further decrease the intensities of already-negligible off-central regions. As a comparison, the intensity of the first-order maxima is approximately $\frac{4}{9\pi^2} \approx 0.0450$ times the intensity of the central maximum. This estimation is obtained from substituting $\alpha = \frac{3\pi}{2}$, which is a rough value and not the true value for the first-order maxima, into Eq. (14.11) (the true value is actually around $\alpha = 4.4934$ radians). In fact, one can show numerically that

[14]A more rigorous approach would use L'Hospital's rule to show that $\lim_{x\to 0} \frac{\sin x}{x} = 1$, which is a well-known result.

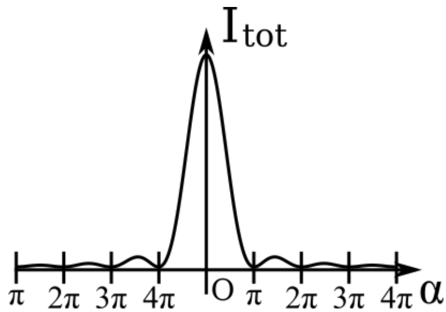


Figure 14.16: Plot of $\operatorname{sinc}^2 \alpha$ against α

roughly 90% of the energy is captured within this central fringe. In light of this, the angular half-width θ of the single slit diffraction pattern is said to correspond to

$$\alpha = \pi \implies \sin\theta = \frac{\lambda}{w}.$$

For small θ,

$$\theta \approx \frac{\lambda}{w},$$

which shows that w must be on the order of λ to produce an observable diffraction pattern. Furthermore, since θ is inversely proportional to w, a narrower slit would produce a wider diffraction pattern. Incidentally, the angular half-width of the diffraction pattern produced by a circular aperture of diameter D is

$$\theta \approx 1.22\frac{\lambda}{D}.$$

The angular half-widths of diffraction patterns have profound ramifications on the resolvability of points in front of a slit. Firstly, observe that if the incident parallel rays on a slit are inclined at an angle α above the horizontal, the entire diffraction pattern simply shifts downwards by an angle α. Then, if the parallel rays from two far-away objects subtend an angle θ, the angular distance between the two peaks of their diffraction patterns is also θ. By the Rayleigh criterion, two objects are said to be resolved if the central maximum of one diffraction pattern lies beyond the first-order minimum of the other. Therefore, the minimum θ between two objects that can be successfully resolved is

$$\theta \approx 1.22\frac{\lambda}{D}$$

for a circular aperture of diameter D, and

$$\theta \approx \frac{\lambda}{w}$$

for a long slit of width w. Finally, note that the presence of a lens (usually spherical) which fills up an aperture does not affect the minimum angular distance of two resolvable objects as the effects of focusing by the lens and diffraction across the aperture can be decoupled and analyzed separately. One can first remove the lens, find the diffraction patterns of the distant objects projected on a screen infinitely far away, behind the slit, and then reflect them about the slit (flip both horizontally and vertically) so that they lie infinitely far away in front of the slit. Subsequently, these reflected diffraction patterns act as the new objects for the lens — the rest is simply a geometrical optics problem. Most notably, the discussion in this paragraph implies that the resolving power R of a telescope with an objective diameter D, defined as the minimum angular distance between two resolvable distant objects, is

$$R \approx 1.22\frac{\lambda}{D},$$

where λ is the observed wavelength of light.

14.1.5 *Interference of Two Wide Slits*

Finally, let us return to a more realistic form of the double slit experiment with slit separation d and a non-negligible slit width w. A plane wave is incident on the slits.

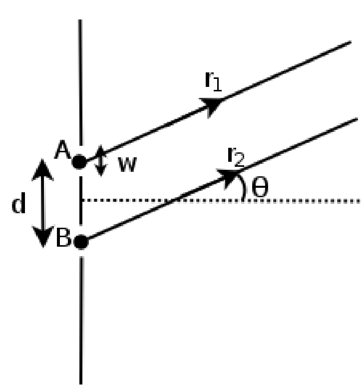


Figure 14.17: Two wide slits

Consider a pair of points, one from each slit, that are a distance d apart. Again, the conditions for constructive and destructive interferences between these points are, respectively,

$$d\sin\theta = n\lambda, \tag{14.12}$$

$$d\sin\theta = \left(n - \frac{1}{2}\right)\lambda. \tag{14.13}$$

This holds for all pairs of points a distance d apart. Evidently, the overall condition for destructive interference due to the two slits is

$$d\sin\theta = \left(n - \frac{1}{2}\right)\lambda.$$

Similarly, without taking into account the phase differences of the waves produced by different pairs of points, the overall condition for constructive interference is

$$d\sin\theta = n\lambda.$$

We are expecting this condition for constructive interference to not be entirely valid as we have not taken into account the phase differences between different pairs. However, the overall condition for destructive interference should be valid as the waves produced by a pair of sources that satisfies the condition annihilate completely — summing these negligible contributions across all pairs should result in zero amplitude and intensity, in spite of the slight phase differences in the contributions across pairs. As anticipated, certain bright fringes that are predicted by the constructive interference condition are not observed empirically.

To take into account the phase difference between pairs and to explain these missing fringes, observe that certain locations on the screen correspond to diffraction minima due to a single slit. Therefore, though the waves from different slits may interfere constructively at a location on the screen, the effect formed on the screen will still be a dark fringe if the diffraction across a single slit results in a diffraction minimum![15] Thus, the missing fringes occur when the condition for the constructive interference between two slits coincides with the condition for destructive interference within a single slit. These requirements are, respectively,

$$d\sin\theta = n\lambda,$$

$$w\sin\theta = m\lambda,$$

[15] $0 + 0$ is still 0.

where $m \neq 0$. Therefore, the nth bright fringe that satisfies

$$n = \frac{dm}{w} \tag{14.14}$$

for some non-zero integer m is missing. Finally, note that we did not really consider the maxima and minima associated with the diffraction across each slit as the intensity plummets from the central fringe of the diffraction pattern such that these off-center maxima and minima become indistinguishable (refer to Fig. 14.16).

Intensity

In this section, we will solve the more general problem of determining the intensity pattern $I_{tot}(\theta)$ for M wide slits of width w and slit separation $d > w$. Define the origin at the center of the bottom-most slit. We have to evaluate the sum of the terms associated with the phases of the waves emanating from each infinitesimal section of all slits. Hence,

$$\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ikr(x)} dx + \int_{-\frac{w}{2}+d}^{\frac{w}{2}+d} e^{ikr(x)} dx + \cdots + \int_{-\frac{w}{2}+(M-1)d}^{\frac{w}{2}+(M-1)d} e^{ikr(x)} dx,$$

which can be rewritten as

$$e^{ikr(0)} \left(\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ikx\sin\theta} dx + \int_{-\frac{w}{2}+d}^{\frac{w}{2}+d} e^{ikx\sin\theta} dx + \cdots + \int_{-\frac{w}{2}+(M-1)d}^{\frac{w}{2}+(M-1)d} e^{ikx\sin\theta} dx \right).$$

The second integral in the brackets is the first integral multiplied by $e^{ikd\sin\theta}$ as its x-coordinates are simply those of the first integral, shifted upwards by a distance d. Concretely, make the substitution $y = x - d$ into the second integral such that it becomes $\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ik(y+d)\sin\theta} dy = e^{ikd\sin\theta} \int_{-\frac{w}{2}}^{\frac{w}{2}} e^{iky\sin\theta} dy$, which makes our claim clear. Similarly, the third integral is $e^{i2kd\sin\theta}$ times the first and so on. The above sum then becomes

$$e^{ikr(0)} \cdot \int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ikx\sin\theta} dx \cdot (1 + r + r^2 + \cdots + r^{M-1}),$$

where $r = e^{ikd\sin\theta}$. Observe that $\int_{-\frac{w}{2}}^{\frac{w}{2}} e^{ikx\sin\theta} dx$ has been determined in the section on a single wide slit and while the geometric series was computed in the section on diffraction gratings. Applying the previous results, the above is

$$e^{i\left(kr(0)+\frac{(M-1)kd\sin\theta}{2}\right)}\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}\cdot\frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{\sin\left(\frac{kd\sin\theta}{2}\right)}.$$

Amalgamating the squared amplitude of the above expression with the $\cos\theta$ term associated with the decay in amplitude of a cylindrical wave,

$$I_{tot}(\theta)\propto\cos\theta\left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{k\sin\theta}{2}}\cdot\frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{\sin\left(\frac{kd\sin\theta}{2}\right)}\right)^2.$$

Following from this, it can be shown that

$$\frac{I_{tot}(\theta)}{I_{tot}(0)}=\cos\theta\left(\frac{\sin\left(\frac{kw\sin\theta}{2}\right)}{\frac{kw\sin\theta}{2}}\cdot\frac{\sin\left(\frac{Mkd\sin\theta}{2}\right)}{M\sin\left(\frac{kd\sin\theta}{2}\right)}\right)^2, \tag{14.15}$$

which is a remarkable result — the intensity ratio of a collection of M wide slits is simply the product of those associated with the diffraction due to a single wide slit and the interference of M thin slits. Since $d \gg w$ under common circumstances, the term associated with diffraction simply acts as a slowly varying envelope (which quickly decays beyond the first diffraction minimum) for the term associated with interference. The observable[16] dark fringes are those associated with the interference term — implying that they occur when $d\sin\theta = (n-\frac{1}{2})\lambda$. Similarly, the observable bright fringes appear when $d\sin\theta = n\lambda$, except for the unusual case where $n = \frac{dm}{w}$ for some non-zero integer m such that the diffraction term yields zero. That said, remember that only an insignificant number of observable bright fringes are missing as the diffraction envelope rapidly decays beyond the first diffraction minimum — implying that the locations, where bright fringes are supposed to appear due to interference but simultaneously correspond to higher-order diffraction minima, are too dim anyway.

14.1.6 *Reflection and Refraction*

Sometimes, waves from a single source (even point sources) can interfere with each other due to reflection and refraction. We shall only be considering light, whose reflection and refraction will be extensively studied later, in our analysis.

Consider the following situation in Fig. 14.18: a monochromatic point source S is placed above an infinite plane mirror. What is the condition for

[16] These are observable precisely because of the contrast with neighboring bright fringes (i.e. they lie within the central diffraction fringe).

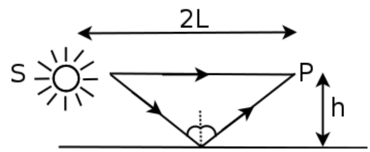


Figure 14.18: Source and mirror

constructive interference to occur at point P, which is at the same height as the source relative to the mirror?

First, we have to identify the waves that interfere at point P. Trivially, one such wave takes a horizontal path from the source to P. Other waves that can possibly reach P are waves that are reflected from the mirror. Since the angle of reflection must be equal to the angle of incidence in the case of reflection, the only path that a reflected beam, which arrives at P, can take is that along the two equal sides of an isosceles triangle, as shown in the diagram above.

A crucial point to note in approaching problems with reflected waves is that there may be a π-radian phase change after reflection. In the case of light waves, if a beam is originally traveling from an optically less dense medium to an optically denser medium, there will be an additional π-radian phase "added" to the reflected wave (besides a change in direction). Otherwise if the beam is traveling from an optically denser medium to an optically less dense medium, the reflected wave will not have an additional π-radian phase change. In this case, the mirror is optically denser than air. Thus, the phase of the reflected wave increases by π radians, which is equivalent to the reflected wave having traveled an additional path length of $\frac{\lambda}{2}$. The condition for constructive interference to occur at point P is then

$$2\sqrt{L^2 + h^2} + \frac{\lambda}{2} - 2L = n\lambda, \qquad (n \in \mathbb{Z}^+)$$

where we have rejected the cases where $n \leq 0$ as the left-hand side is always greater than zero. If $h \ll L$, a slick analysis exists. Consider the mirror image S' of the source that is a distance h below the mirror (Fig. 14.19). The path distance traveled by a reflected ray from the real source S to an arbitrary point P is equal to a ray that emanates from the virtual source S' and directly travels to P. However, S' is effectively perfectly out of phase relative to S due to the π-radian phase shift at the mirror. Then, we can apply a procedure similar to the double slit experiment with the slit separation $d = 2h$ to determine the condition for constructive interference at any point P which is a horizontal distance $2L$ away from the source (not necessarily aligned with S now).

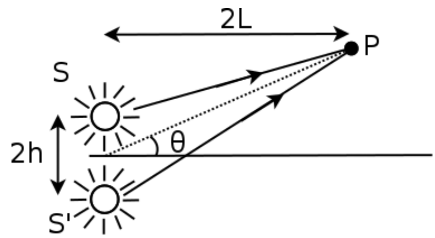


Figure 14.19: Source S and virtual source S'

If we define θ to be the angle between the intersection of the mirror with the line joining the two sources and point P,[17] the condition for constructive interference is

$$2h\sin\theta = \left(n - \frac{1}{2}\right)\lambda \qquad (n \in \mathbb{Z}^+)$$

where the $-\frac{1}{2}\lambda$ accounts for the virtual source being out of phase with the real source. One can show that this is consistent with the previous equation for $\sin\theta = \frac{h}{2l}$ (point P is on the same vertical level as S as $\tan\theta \approx \sin\theta$ for small θ), after performing a binomial expansion for $\sqrt{L^2 + h^2}$ and discarding higher-order terms in $\frac{h^2}{L^2}$.

Bragg's Law

It was discovered that crystalline solids produced intensity maxima and minima of reflected radiation when illuminated by X-rays of specific wavelengths and incident angles. A model was proposed to explain this strange phenomena — a crystal is imagined to be a set of myriad, discrete and one-atom-thick planes of atoms separated by a uniform distance d.

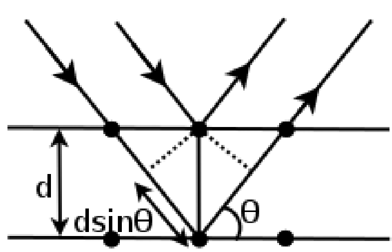


Figure 14.20: Layers of atoms

[17]If $h \ll L$, θ should really also be the angle between the emitted rays to P and the horizontal.

The incident EM waves are scattered by the layers of atoms, indicated by black circles in Fig. 14.20. Then, the Bragg peaks were proposed to occur at positions where the reflected beams from these atomic layers interfere constructively (the rays are not exactly parallel but intersect far away as the incident rays are also not exactly parallel if they came from the same source). As seen from the diagram on the previous page, the path difference of rays between two adjacent layers is

$$\delta = 2d\sin\theta.$$

Even though there is a phase change of π radians during reflection, it occurs for all of the reflected waves. Hence, the condition for constructive interference is still

$$\delta = 2d\sin\theta = n\lambda, \qquad (n \in \mathbb{Z}^+)$$

while the condition for destructive interference is

$$\delta = 2d\sin\theta = \left(n - \frac{1}{2}\right)\lambda. \qquad (n \in \mathbb{Z}^+)$$

Thin Film Interference

When white light is shone on a thin soap film, it appears iridescent. The cascade of colours is a result of the constructive and destructive interferences between certain reflected waves from the two surfaces (a soap film has a top and a bottom surface). Thus, some reflected colours are "canceled" out and the remaining colors are observed — certain colors may even be amplified. Consider the following diagram.

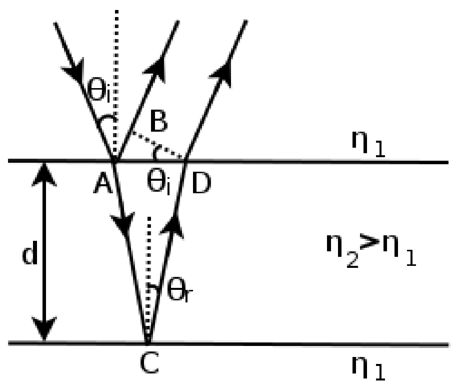


Figure 14.21: Thin film

Monochromatic light of wavelength λ in vacuum and frequency f travels from an optically less dense medium, with an absolute refractive index n_1, to an optically denser medium, of an absolute refractive index n_2, at an angle of incidence θ_i. Part of the light is transmitted and part of it is reflected at the top surface. The transmitted light wave then propagates in the medium before being partially reflected at the bottom surface and being partially transmitted into the less dense medium again — the reflected rays then interfere (though they are parallel, they are usually focused at a single point by an optical apparatus, such as the lenses in our eyes). In this analysis, it is important to note that light slows down in an optically denser medium. Furthermore, since

$$c = f\lambda$$

in vacuum and the phase velocity of light in a denser medium is $c' = \frac{c}{n}$, where n is the absolute refractive index of the medium, the wavelength of light in a medium of refractive index n must be

$$\lambda' = \frac{\lambda}{n},$$

as the frequency is an intrinsic property of the wave source and should not vary when transiting across a stationary interface. This means that a certain amount of path length in the optically denser medium corresponds to an effective phase change that is larger by a factor of n as compared to the corresponding phase change if the wave were to travel the same path length in vacuum. Therefore, it is useful to define a quantity known as the optical path length (OPL) in a medium, which is given by

$$OPL = nl,$$

where l is the path length traversed by the light in that medium. Then, the relationship between the optical path length covered and the phase change is

$$\frac{\Delta\phi}{2\pi} = \frac{OPL}{\lambda},$$

where λ is the wavelength of the light in vacuum or air. In light of this, the optical path difference between the two emerging reflected rays is

$$\delta_{OPL} = n_2\left(\overline{AC} + \overline{CD}\right) - n_1\overline{AB}.$$

The length of path ADC is given by

$$\overline{AC} + \overline{CD} = \frac{2d}{\cos\theta_r},$$

where d is the thickness of the soap film. The length of AD is

$$\overline{AD} = 2d\tan\theta_r.$$

The length of AB is then

$$\overline{AB} = \overline{AD}\sin\theta_i = 2d\tan\theta_r\sin\theta_i.$$

Furthermore, by applying Snell's law,

$$n_1\sin\theta_i = n_2\sin\theta_r,$$

the optical path difference can be expressed as

$$\delta_{OPL} = 2n_2d\left(\frac{1}{\cos\theta_r} - \frac{\sin^2\theta_r}{\cos\theta_r}\right) = 2n_2d\cos\theta_r.$$

Lastly, since the reflected wave at the top undergoes a π-radian phase shift while that at the bottom does not, the condition for constructive interference between the two emerging reflected waves is

$$2n_2d\cos\theta_r = \left(n - \frac{1}{2}\right)\lambda. \qquad (n \in \mathbb{Z}^+)$$

Remember that λ refers to the wavelength of the light in a vacuum. Next, the condition for destructive interference is simply

$$2n_2d\cos\theta_r = n\lambda. \qquad (n \in \mathbb{Z}^+)$$

Alternatively, these conditions can be expressed in terms of n_1 and θ_i but the above form looks neater.

Problem: If an infinitesimally thin soap film is illuminated by white light normal to its surface, what do you expect to observe? The two surfaces of the film form interfaces with air.

You would expect to observe a black soap film. This is because the optical path difference is insignificant, in this case. Only the phase shift of π radians matters, resulting in destructive interference between pairs of reflected rays of all wavelengths.

14.2 Standing Waves

Consider two coincident one-dimensional waves with displacement functions

$$\psi_1(x,t) = A\sin\left(kx - \omega t + (\phi_1 - \phi_2)\right),$$
$$\psi_2(x,t) = A\sin\left(kx + \omega t + (\phi_1 + \phi_2)\right).$$

These equations represent two traveling waves of equal amplitude, speed and wavelength but opposite directions of travel. The phase offsets look rather ugly now, but they are expressed as such for future simplicity. Applying the trigonometric identity $\sin A + \sin B = 2 \sin \frac{A+B}{2} \cos \frac{A-B}{2}$, the resultant displacement due to these waves is

$$\psi(x,t) = A \sin(kx - \omega t + (\phi_1 - \phi_2)) + A \sin(kx + \omega t + (\phi_1 + \phi_2)) \tag{14.16}$$

$$= 2A \sin(kx + \phi_1) \cos(\omega t + \phi_2). \tag{14.17}$$

From this expression, we see that each point on the resultant wave oscillates about its equilibrium position with an amplitude $|2 \sin(kx + \phi_1)|$ and an angular frequency ω, where x is the x-coordinate of that point. ϕ_2 is a constant that only depends on the choice of the origin of time and can be tweaked to be whatever we want. On another note, the time-component of Eq. (14.17) is largely irrelevant in our analysis, which will focus on the amplitude of oscillation.

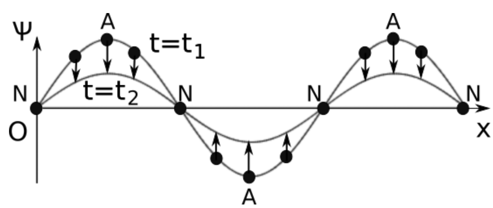


Figure 14.22: Resultant displacements at two instances

The graph above shows the displacement against position of the points on the resultant wave at two different times, t_1 and t_2 (with $\phi_1 = 0$). Evidently, the wave profile does not advance with time (it is merely squashed or enlarged with time), so no net energy is transferred from one point in space to another, on average.[18]

Furthermore, there are some locations at which the points on the resultant wave do not oscillate at all. These correspond to points known as nodes whose x-coordinates satisfy $x = \frac{n\lambda}{2}$, $n \in \mathbb{Z}$ in this case. On the other hand, points on the wave which oscillate at the maximum amplitude $2A$ are known as antinodes. In this case, the x-coordinates of the antinodes fulfil $x = (\frac{n}{2} - \frac{1}{4})\lambda$, $n \in \mathbb{Z}$.

There is always an antinode in the middle of two nodes and vice-versa. Half of the common wavelength of the two original waves, $\frac{\lambda}{2}$, is equal to the distance between two adjacent antinodes or nodes. Lastly, the phase of a

[18] We have shown in Problem 4 (Chapter 13) that the power transmitted across a point due to the superposition of two waves traveling in opposite directions is simply the sum of the individual powers. Taking the time-average would yield zero. However, note that there may still be instantaneous net power transmitted across a point.

particular point on the resultant wave depends on the sign of $\sin(kx + \phi_1)$. If $\sin(kx + \phi_1)$ is positive, the phase of that point is $(\omega t + \phi_2)$. Otherwise, it is $(\omega t + \phi_2 + \pi)$, as an additional π radians in phase angle would negate the instantaneous displacement of a point. Since the sine function is positive and negative for π radians each in a single period, and the distance between consecutive nodes is $\frac{\lambda}{2}$, the phases of all points between 2 adjacent nodes are the same while the phases of points in adjacent segments between nodes differ by π radians.

The resultant wave above is known as a standing wave and is formed when two identical waves — of the same amplitude, frequency and speed but opposite directions of travel — are superposed. Standing waves are pervasive and play an essential role in musical instruments, as we shall see. In general, waves are not necessarily sinusoidal but a general wave can be expressed as an algebraic sum of an infinite series of sinusoidal waves via Fourier analysis (thus, sinusoidal waves are useful). Different linear combinations of the sinusoidal standing waves in an instrument produce the unique sounds that we hear. Finally, the wavelengths and thus frequencies of standing waves in an object are actually restricted by its physical parameters to a certain set of values — explaining why many instruments can be tuned, by varying its physical characteristics, to produce a certain pitch. This will be explored in the next section.

14.2.1 *Boundary Conditions*

Now, we will analyze the sinusoidal solutions in certain set-ups which set a foundation for more general solutions. For a single sinusoid, the value of ϕ_1 in Eq. (14.17) is determined by a single boundary condition such as the location of a node or antinode. Additional boundary conditions then impose constraints on the possible values of k, as we shall see.

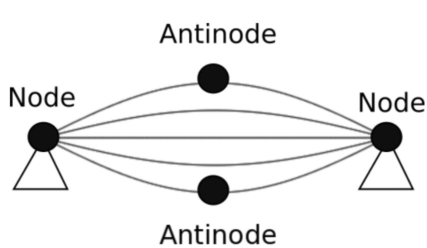


Figure 14.23: Guitar string with clamps

Consider a uniform guitar string that is clamped on both ends a distance L apart and plucked. Transverse waves travel along the string and are completely reflected from both ends to produce a stationary wave (hard reflection). Figure 14.23 shows the resultant displacement of the string at different times if the string was initially plucked in the middle to form a sinusoidal wave profile.

Regardless of how the string was actually disturbed initially, the two ends of the string must remain stationary and must thus be displacement nodes. If we let the origin of the x-axis be at the left end of the string, the condition for the left end to be fixed at all times for a single sinusoid can be obtained from Eq. (14.17). Since the amplitude of oscillation at a coordinate x is $|2A\sin(kx+\phi_1)|$, the node condition for the left end at which we set $x=0$ is satisfied if

$$\sin\phi_1 = 0.$$

Whether $\phi_1 = 0$ or π does not matter as we are only concerned about the amplitude of oscillation. A π-phase shift only introduces an additional negative sign. Therefore, we choose $\phi_1 = 0$ and substitute it into Eq. (14.17) to get

$$\psi(x,t) = 2A\sin kx\cos(\omega t+\phi_2).$$

Since the point at $x = L$ must also be a displacement node,

$$\sin kL = 0$$
$$\implies k = \frac{n\pi}{L}, \qquad (n \in \mathbb{Z}^+)$$
$$\lambda = \frac{2\pi}{k} = \frac{2L}{n}.$$

Let v be the speed of the string which is independent of the boundary conditions, as $v = \sqrt{\frac{T}{\mu}}$ where T is the tension in the string and μ is the linear mass density. Thus, from $v = f\lambda$,

$$f = \frac{nv}{2L}.$$

The above equation implies that the frequencies produced by a string clamped on both ends can be changed by tightening the clamps (to vary the tension and thus vary v) or changing the length of the string L. Moving on, each value of n corresponds to a different mode of oscillation of the string with a unique wavelength as shown in Fig. 14.24, for which the string has an

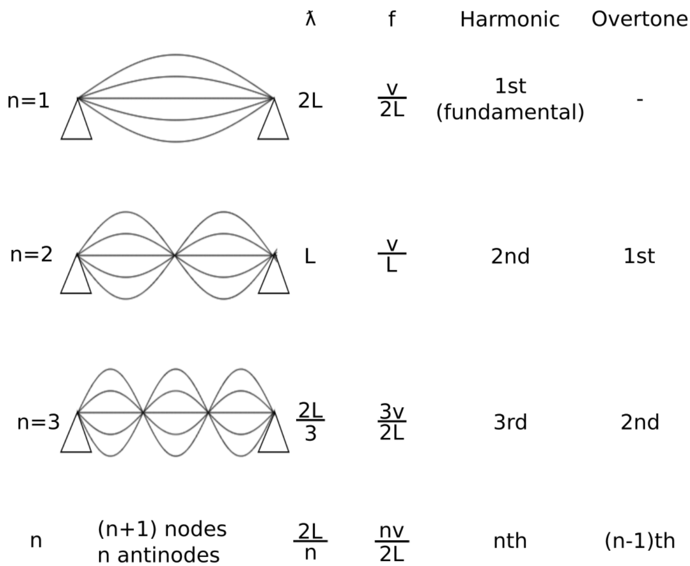


Figure 14.24: Guitar string with clamps

infinite number.[19] The figure above depicts the standing waves at different instances — however, for all diagrams in the future, we will only draw the amplitude envelope for the sake of brevity.

All possible standing wave configurations on the string in this case have two nodes at its ends and any number of nodes distributed evenly between its ends. The midpoint between every two successive nodes corresponds to an antinode. Evidently, since the wavelengths of the standing waves can only take on certain values, the pure frequencies that can be produced by the string also take on certain discrete values. These are known as the **resonant frequencies** or **natural frequencies** of the system, aptly termed as such because the system will respond with the largest amplitude at a particular resonant frequency when the frequency of the external source matches that particular resonant frequency. Otherwise, a negligible response will be obtained.

The fundamental frequency, f_1, corresponds to the lowest possible resonant frequency of a system. In this case, $f = \frac{v}{2L}$. The nth ($n \in \mathbb{Z}^+$) harmonic refers to the frequency that results from a certain mode of oscillation that is n times the fundamental frequency, i.e. $f_n = nf_1$. In this particular case, all harmonics are present, but this may not be true in general.

[19] This is to be expected, as we can consider each segment to be an oscillator. The continuous string is then an infinite array of coupled oscillators which has an infinite number of modes.

Overtones are the possible resonant frequencies that can be produced by a system — excluding the fundamental frequency. Again, they are sorted and numbered in ascending order. In this case, the first overtone of the string corresponds to the 2^{nd} harmonic.

Sound Waves in Pipes

When air is blown into a pipe, a certain sound is heard. Furthermore, the sound seems to differ across different pipes. A familiar example would be acoustic performances where artists blow into beakers containing different levels of water to produce different tunes. This effect is due to the sound being a mixture of certain attainable frequencies which are dependent on the physical parameters of the tube.

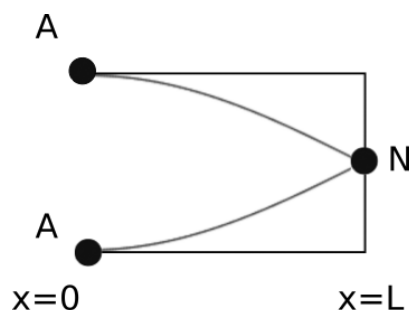


Figure 14.25: Displacement of air molecules

The graph in the diagram above shows the amplitudes of oscillation of the air molecules in a pipe that is closed at one end, due to a possible sinusoidal standing wave at a certain instance in time. Note that even though the amplitude envelope is depicted in the transverse direction for purposes of illustration, the air molecules are actually displaced in the longitudinal direction.

At the closed end of the pipe, the air molecules can neither penetrate the wall nor escape from it as the resulting vacuum would pull it back — implying that they are unable to oscillate. Thus, the closed end must correspond to a displacement node. The displacement wave undergoes a π-radian phase change after reflection and "cancels" the incident wave. Another reason would be that the pressure wave undergoes no phase change after reflection from the wall as the closed end has a higher acoustic impedance.[20] Thus,

[20]The reader should convince himself or herself that when the displacement wave undergoes a π-radian phase change upon reflection, the pressure wave undergoes no phase change

the reflected pressure wave will interfere constructively with the incident pressure wave at the closed end of the pipe — causing the closed end of the pipe to correspond to a pressure antinode and hence, a displacement node.

Most intuitively, the air molecules at the open end should be free to move — causing a displacement antinode to be located there. A more rigorous explanation is that the pressure of air at the open end of the tube does not deviate much from the external air pressure due to the enormous volume of air in the surroundings. We do not expect to produce a significant change in its pressure through actions such as blowing into the tube. Thus, the point at the open end must be a pressure node. A better reason would pertain to the phase shift of π radians of the pressure waves that are reflected at the open end of the pipe,[21] which causes them to interfere destructively with the incident waves at the open end — producing a pressure node at the open end and thus a displacement antinode. Definitely, the incident pressure wave is not reflected completely and partially escapes — evident from the fact that sound is still heard which implies energy transfer into the surroundings. However, a displacement antinode at the open end is still a decent approximation in this situation.

Define the origin to be located at the closed end of the tube and the x-coordinate of the open end to be $x = -L$. Again, $\phi_1 = 0$ in Eq. (14.17) for the displacement standing wave to ensure that the closed end corresponds to a displacement node. Next, in order for $x = -L$ to be a displacement antinode,

$$|\sin kL| = 1$$

$$kL = \left(n - \frac{1}{2}\right)\pi \qquad (n \in \mathbb{Z}^+)$$

$$k = \frac{2n-1}{2L}\pi,$$

$$\lambda = \frac{4}{2n-1}L$$

upon reflection. Keep in mind that the pressure and displacement waves are $\frac{\pi}{2}$ out of phase. This phase difference is amplified by a factor of two when the displacement wave reflects, which leads to there being no phase change in the pressure wave.

[21] Yes, sound waves are partially reflected at the open end of the pipe as it travels into a region of a lower impedance (as the air outside the pipe is generally less dense) — akin to how light is partially reflected when incident upon a medium with a lower refractive index. The reflected displacement wave does not experience a phase shift — causing the reflected pressure wave to undergo a π-radian phase shift.

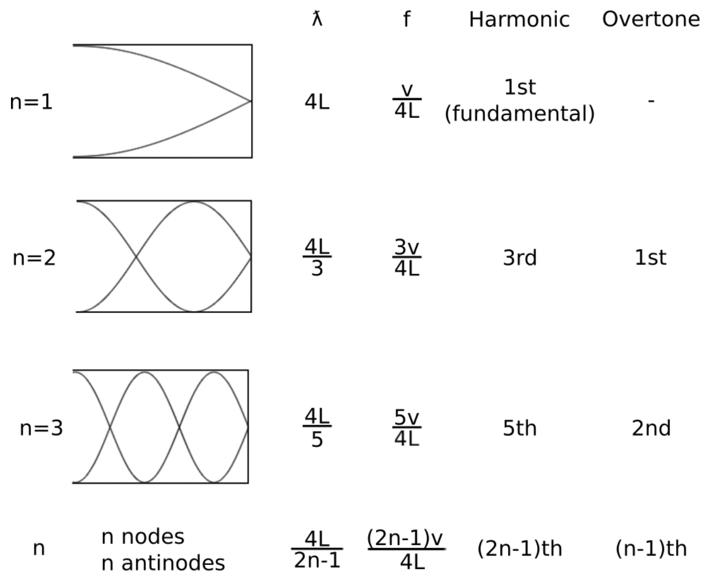


Figure 14.26: Harmonics and overtones

These are the possible wave numbers and wavelengths of the sinusoidal standing waves inside the pipe. Figure 14.26 depicts the modes of oscillations for $n = 1, 2, 3$.

As seen from above, there can be any number of displacement nodes and antinodes between the open and closed ends of the pipe, which correspond to a displacement antinode and node respectively. The standing wave, with the mode of oscillation conforming to a value n, has n antinodes and nodes each.

The fundamental frequency in this case is

$$f_1 = \frac{v}{4L}.$$

It is important to observe that only the odd harmonics are present in a tube with one closed and one open end. An example of such an instrument would be the flute. To play a flute, one blows into the side of the flute near the top, which is similar to an open end, while the other end of the flute is blocked and is hence a closed end.

Similarly, for a pipe of length L that is open on both ends, the two ends are displacement antinodes and the standing waves can only take on specific wavelengths too. To determine the attainable wavelengths, consider the fact that the amplitude of the displacement wave at a certain point with coordinate x is given by Eq. (14.17) as $|2A \sin(kx + \phi_1)|$. Therefore, if we choose the origin at the left end of the pipe, the displacement function must satisfy the following condition for the open end at $x = 0$ to be an antinode:

$$\sin \phi_1 = 1,$$

which implies that $\phi_1 = \frac{\pi}{2}$ and $|2A \sin(kx + \phi_1)| = |2A \sin(kx + \frac{\pi}{2})| = |2A \sin(\frac{\pi}{2} - kx)| = |2A \cos kx|$.

For the end at $x = L$ to also be an antinode,

$$|\cos kL| = 1$$

$$k = \frac{n\pi}{L}, \qquad (n \in \mathbb{Z}^+)$$

$$\lambda = \frac{2L}{n}.$$

Figure 14.27 shows the different displacement modes of oscillation for $n = 1, 2, 3$.

Contrary to a pipe with a single closed end, an open pipe can produce all possible harmonics. Once again, there can be any number of nodes in between the two ends as long as the two ends remain as antinodes. In this case, there must at least be one node between the two ends as there must be a minimum of one node between two antinodes.

Ultimately, we do not usually consider the equations of standing waves when solving for the possible modes of oscillation, the way we did in this

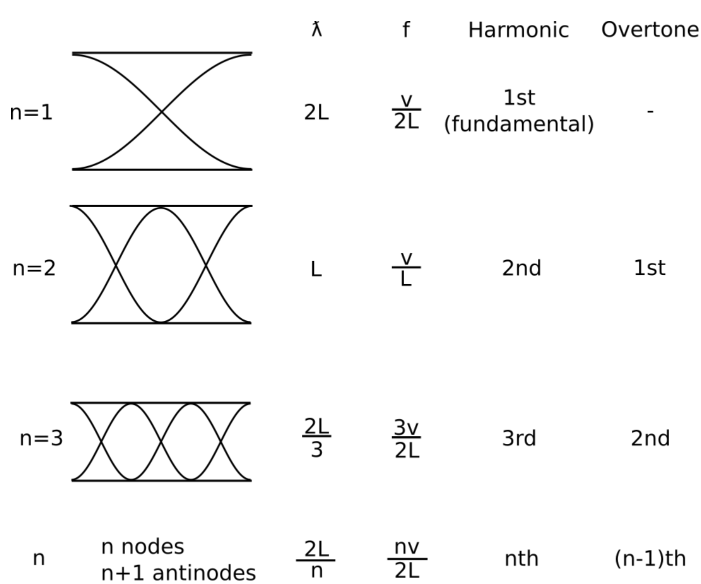


Figure 14.27: Harmonics and overtones

section. Instead, drawing the possible displacement waves in a system, while considering whether the boundaries are constrained to be nodes or antinodes, is often more intuitive, edifying and efficient. Thus, this method should be adopted in most problem-solving practices.

A Comment

In reality, an antinode is not situated perfectly at the open end of a pipe. In fact, it occurs slightly outside of a pipe. This is known as the end correction for open pipes which is experimentally determined to be of distance $0.6r$ for a cylindrical pipe, where r is the radius of the pipe. Thus, it is a property of the pipe itself and is independent of the frequency of the standing wave. When solving problems, it is paramount to choose methods that can account for such end corrections if possible. Consider the following example.

An air column in a cylindrical glass pipe is blocked by a movable piston on one end and is open on the other end. A tuning fork of frequency f is placed at the open end of the tube. The movable piston is then slowly extracted to increase the distance between the closed and open ends. Resonance is first heard when the piston is of the minimum distance d_1 away from the open end, and next heard when the piston is distance d_2 away from the open end. Determine the speed of sound using these parameters.

Well, we might say that the situation when resonance is first heard corresponds to the fundamental mode of oscillation and conclude that

$$\lambda = 4d_1$$
$$\implies v = f\lambda = 4fd_1.$$

However, this does not account for possible end effects which may be significant if the radius of the pipe is large. That is, $d_1 \neq \frac{\lambda}{4}$ but rather, $d_1 = \frac{\lambda}{4} - c$ where c is the end correction term that is independent of the mode of oscillation. Instead, it is more accurate to consider the difference between the distances in the two situations. During the second time resonance is heard, there is now an additional displacement node in the middle of the two ends as compared to the first situation. Thus, this must correspond to an increase of $\frac{\lambda}{2}$ in the length of the pipe.

$$d_2 - d_1 = \frac{\lambda}{2},$$
$$v = f\lambda = 2f(d_2 - d_1),$$

where we have accounted for possible end effects.

14.3 Beats

Another application of the principle of superposition in music is the overlap of two sinusoidal traveling waves of slightly different frequencies. Then, a periodic "beat" or a fluctuation in the intensity of the sound wave can be heard. Quantitatively, consider two waves at a certain point in space which result in individual displacements at that point of the form

$$\psi_1 = A\cos(2\pi f_1 t + \phi_1),$$
$$\psi_2 = A\cos(2\pi f_2 t + \phi_2).$$

Superposing these waves, we obtain the resultant displacement at that point ψ_R.

$$\begin{aligned}\psi_R &= \psi_1 + \psi_2 \\ &= A[\cos(2\pi f_1 t + \phi_1) + \cos(2\pi f_2 t + \phi_2)] \\ &= 2A\cos\left(2\pi\frac{f_1+f_2}{2}t + \frac{\phi_1+\phi_2}{2}\right)\cos\left(2\pi\frac{f_1-f_2}{2}t + \frac{\phi_1-\phi_2}{2}\right).\end{aligned}$$

Scrutinizing this expression, if $f_1 \approx f_2$, $A\cos\left(2\pi\frac{f_1-f_2}{2}t + \frac{\phi_1-\phi_2}{2}\right)$ can be treated as a slowly varying amplitude which acts as an envelope for the other oscillatory term. The resultant displacement at that point then seemingly oscillates with a frequency $\frac{f_1+f_2}{2}$ under the envelope. Plotting this function on a graph, we expect a large sinusoidal envelope with rapid sinusoidal fluctuations within the envelope.

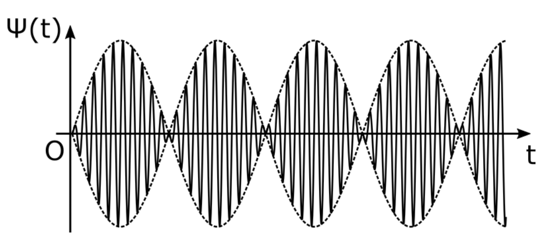


Figure 14.28: Graph of resultant displacement

In the graph above, the horizontal axis indicates time while the vertical axis indicates the instantaneous displacement of the resultant wave at that particular location. As observed from the graph above, the envelope of the graph traces out a sinusoidal function with a large period and encapsulates rapid fluctuations.

There are global intensity maxima (where the displacement is largest in absolute value) and minima (where the displacement is zero) on the envelope. A beat is heard when there is a transition from one global maximum to another. Thus, the beat frequency is $f_{beat} = |f_1 - f_2|$ rather than $\frac{|f_1 - f_2|}{2}$ which is the frequency of the slowly oscillating cosine function. This is due to the fact that one complete cycle of oscillation of the envelope corresponds to two intensity jumps (maximum positive displacement to maximum negative displacement and back to maximum positive displacement) and thus two beats.

Musical beats are pivotal in tuning instruments. One can vibrate a tuning fork and a piano string at the same time. If their frequencies are out of sync, a periodic beat will be heard. Then, one can adjust the tension in the piano string so that the frequency of beats decreases until they eventually become virtually undetectable. Then, the string is tuned.

Experiment: Go to http://onlinetonegenerator.com and play a sinusoidal sound at 440 Hertz. With the current tab open, open a duplicate tab and play a sound at 441 Hertz to hear a rhythmic beat. Now, increase the frequency of the sound from 441 Hertz to 442 Hertz and empirically verify that the frequency of the beat increases.

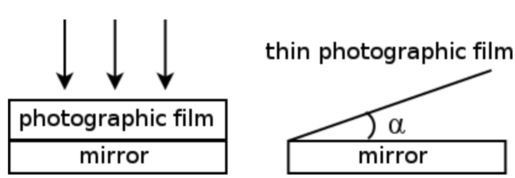
Problems

1. *Double Standing Waves**

Two in-phase monochromatic lasers of wavelength λ are located at $(-l_1, 0)$ and $(0, -l_2)$ in the xy-plane. They simultaneously fire horizontal and vertical beams respectively which are then reflected by perfectly reflective mirrors at $(l_1, 0)$ and $(0, l_2)$, whose planes are perpendicular to the incoming rays — setting up two standing waves. The directions of oscillation of the standing waves are aligned. If the two antinodes of the standing waves nearest to the mirrors attain their maximum displacements (in the same direction) simultaneously at all times, determine the respective conditions for the intensity at the origin to be larger and smaller than the individual intensity caused by either standing wave.

2. *Ives' and Wiener's Experiments**

The left and right set-ups are used by Ives and Wiener to measure the wavelength λ of incident light respectively. Monochromatic light is normally incident on the two mirrors in the two experiments. If the photographic film is transparent and can capture regions of different intensities, how can λ be determined? Form an equation for λ in terms of the measurable quantities on the photographic films and angle α in the case of Wiener's experiment. In Wiener's set-up, the photographic film is inclined by a small angle $\alpha \ll 1$; what is the advantage of this?



3. *Displaced Slit**

Consider Fig. 14.8. If the light source S and slit S_0 are both shifted downwards so that they are aligned with the bottom slit, quantitatively describe the changes to the locations of the bright and dark fringes for small θ (the definition for θ is retained). Let the horizontal distance between S and S_0 be w such that $\lambda \ll w$ but the squared slit separation d^2 is comparable with λw. $L \gg d$ is the horizontal distance between the double slit and the screen.

4. *Glass Plate**

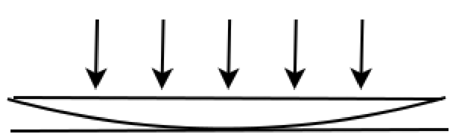
A thin glass plate of refractive index n and thickness t is placed directly in front of the top slit of a double slit experiment with slit separation d and wavelength λ in vacuum. Quantitatively describe the resultant change in the locations of the bright and dark fringes on a screen located far away at a distance L, such that $L \gg t$ and $L \gg d$, for small values of angle θ (for which the conventional definition is adopted).

5. *Wedge**

A glass wedge with an unknown small wedge angle α is suspended with its tip at the top. Consider a uniform triangular cross-section of the wedge. A movable source emitting a narrow, horizontal beam of monochromatic light is then shone from the left side of the wedge, with a progressively increasing vertical distance y from the tip of the edge. At a small distance $y = y_2$, the second-order intensity maximum is observed from the left of the wedge. At what distance y_3' will the third-order intensity minimum be observed?

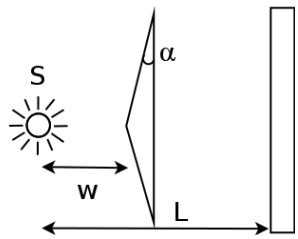
6. *Newton's Rings***

A thin plano-convex lens with a large radius of curvature R lies on the surface of a table as shown below. There is a thin air gap between the surface of the lens and the table, at points other than the vertex of the lens. Monochromatic light of wavelength λ is then shone normally on the lens (vertically downwards) and a series of bright and dark circles are observed when the lens is viewed from above. Determine the radii of the nth bright and dark rings.



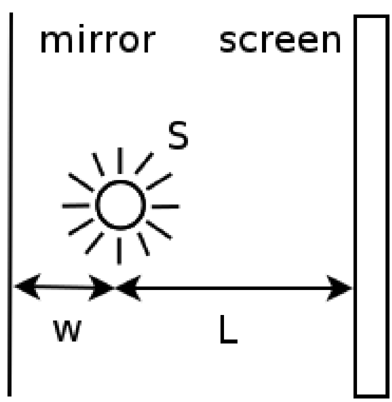
7. *Fresnel's Biprism***

A monochromatic cylindrical source S of wavelength λ is placed a distance w in front of a thin, infinitely long biprism (two prisms glued together to form an isosceles triangle) of small angle α. The refractive index of the biprism is n. A screen is placed far away from the source S and an interference pattern is observed. By considering the notion of virtual sources, explain why an interference pattern is formed despite there being only a sole point source. Then, let θ denote the angle subtended by a line joining S to a point on the screen and the horizontal. Determine the θ coordinate of the mth-order bright fringe. Finally, each virtual source has a limited angular range of illumination. Determine the total number of bright fringes formed on the screen while assuming $w \ll L$, where L is the distance between S and the screen.



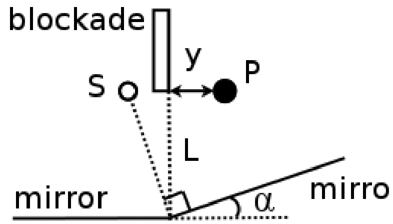
8. *Mirror Interference***

A monochromatic point source S of wavelength λ is placed at a distance $w = k\lambda$ away from an infinite vertical mirror. A screen is then placed a distance $L \gg w$ from S on the other side. Determine the total number of bright rings produced by this set-up.



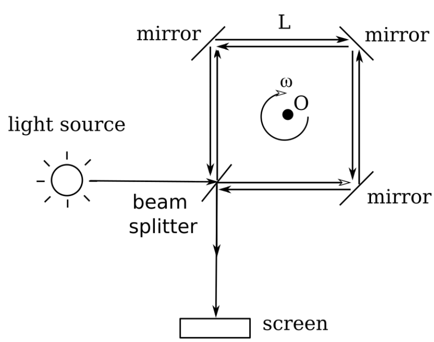
9. *Slanted Mirrors***

A monochromatic point source S of wavelength λ and two mirrors are arranged as shown in the set-up on the next page (where α is small). S is a vertical distance L above the horizontal mirror. A thin blockade is placed on the left on the middle line to prevent horizontal rays from S from passing through. The bottom of the blockade is aligned with S. Determine the condition for constructive interference at a point P, which is at the same vertical level as S and a horizontal distance y rightwards from the center line. Determine the total number of possible points P (with different y) at which constructive interference can occur.



10. ***Sagnac Interferometer*****

Three mirrors and a beam splitter are arranged into a square of side length L as shown in the figure below. A sinusoidal light source of period T first emits a ray into the beam splitter which splits the incident ray into two that travel upwards and rightwards respectively. These rays then travel one round along the edges of the square before being recombined again by the beam splitter. Suppose that the mirrors rotate about the center of the square O with angular frequency ω ($\omega L \ll c$ where c is the speed of light), the two recombined rays will have a phase difference. If L is the smallest length for which constructive interference occurs on the screen, determine ω.



11. ***Accelerating Cars*****

Two cars are initially located on the x-axis at $x = -l$ and $x = l$ respectively. The sirens on the cars are in-phase and emit high-frequency sound waves of small period T, in the frame of the sirens, along the x-axis. The sound waves propagate at speed c in air. For $t \geq 0$, the cars travel towards each other at speed $v(t) = v_0 - at$ where $v_0 > c$ is the initial velocity and a is a constant acceleration. At an unknown time t', a stationary observer at $x = d$ where $0 < d < l$ receives waves from the two sources that have zero phase difference. Find t'. Then, determine the conditions for the observer to only receive one wave from each source at time t', as well as the times at which the cars emitted the waves that reached the observer at time t', under these conditions.

Solutions

1. Double Standing Waves*

Due to their perfectly reflective properties, the mirrors produce nodes as hard reflection causes the reflected wave to annihilate the incident ray at the point of reflection. Next, since the antinodes of the two standing waves closest to the mirrors attain their maximum displacements simultaneously, the two time-dependent terms (refer to Eq. (14.17)) of the two standing waves must be identical. Therefore, we can simply consider their amplitudes in determining the resultant intensity.

We know that two adjacent segments between nodes in a single standing are π radians apart in phase. Therefore, the point at the origin is either an odd or even multiple of π radians apart in phase from the segment containing the mirror node for each standing wave. We simply need to determine whether the segments of the two standing waves that contain the origin are an even or odd multiple of π radians in phase apart to determine if the combined intensity is larger or smaller than either of the individual intensities — an even multiple leads to a larger resultant intensity while the converse holds as well. Since the nodes of a standing wave are $\frac{\lambda}{2}$ apart, for the horizontal wave, the segment containing the origin is the $\left\lceil \frac{2l_1}{\lambda} \right\rceil$th segment from the mirror. A similar statement holds for the vertical wave. Therefore, for the two segments at the origin to be an even multiple of π radians apart and the resultant intensity to be larger than the individual intensities,

$$2 \mid \left\lceil \frac{2l_1}{\lambda} \right\rceil + \left\lceil \frac{2l_2}{\lambda} \right\rceil.$$

Otherwise, for the resultant intensity to be smaller,

$$2 \nmid \left\lceil \frac{2l_1}{\lambda} \right\rceil + \left\lceil \frac{2l_2}{\lambda} \right\rceil.$$

Note that if either standing wave forms a node at the origin ($\frac{2l_1}{\lambda}$ or $\frac{2l_2}{\lambda}$ is an integer), the resultant intensity will equal to the individual intensity of the other wave. Thus, we must exclude such possibilities.

2. Ives' and Wiener's Experiments*

The incident ray is reflected by the mirror which sets up a vertical standing wave with a node located at the mirror (due to the π-radian phase shift owing to hard reflection). The photographic film then records the regions of high intensity and low intensity, which correspond to antinodes and nodes respectively. Since we know that the vertical distance between adjacent antinodes must be $\frac{\lambda}{2}$ apart, we can determine λ by measuring the distance between adjacent high intensity regions on the photographic film, Δd (we can also compute the total distance between many high intensity regions and take the average to reduce the percentage error of measurement). For Ives' set-up,

$$\lambda = 2\Delta d.$$

For Wiener's set-up, we have to take the inclination into account. The distance between adjacent high intensity regions on the photographic film is

$$\Delta d = \frac{\lambda}{2\sin\alpha} \approx \frac{\lambda}{2\alpha}.$$

Note that we divide by $\sin\alpha$ to obtain the "hypotenuse" from the "adjacent side". The advantage of having a small angle of inclination is that Δd is greatly amplified such that the percentage error in measuring it is reduced. λ can then be expressed in terms of Δd and α as

$$\lambda = 2\alpha\Delta d.$$

3. Displaced Slit*

Due to the displaced slit S_0, the rays that travel to the top slit cover a longer path length than those that travel to the bottom slit — this additional path length is $\sqrt{w^2+d^2} - w$. Therefore, the path difference between rays that emanate from S and end at a point P on the screen through the two slits is $d\sin\theta - \sqrt{w^2+d^2} + w$. For constructive interference,

$$d\sin\theta - \sqrt{w^2+d^2} + w = n\lambda.$$

We can perform a binomial expansion for the square root term:

$$\begin{aligned}\sqrt{w^2+d^2} &= w\sqrt{1+\frac{d^2}{w^2}} \\ &= w\left(1+\frac{d^2}{2w^2} - \frac{d^4}{8w^4} + \cdots\right) \\ &= w + \frac{d^2}{2w} - \frac{d^4}{8w^3} + \cdots \\ &\approx w + \frac{d^2}{2w}.\end{aligned}$$

Since we will be comparing the terms with λ, we note that since d^2 is comparable with $w\lambda$, $\frac{d^2}{w}$ is comparable with λ, and we keep the second term.

However, we do not keep higher-order terms, such as $\frac{d^4}{w^3} \sim \frac{w^2\lambda^2}{w^3} = \frac{\lambda^2}{w}$, which is negligible when compared to λ as $\lambda \ll w$. Then, the condition for constructive interference is

$$d \sin\theta = n\lambda + \frac{d^2}{2w}.$$

For small θ, $\sin\theta \approx \frac{y}{L}$ where y is the vertical coordinate of the point on the screen at which the rays converge, above the center of the two slits. Then,

$$y = \frac{n\lambda L}{d} + \frac{dL}{2w}.$$

By comparing this with the location of bright fringes in a normal double slit experiment, it can be seen that the bright fringes are displaced by $\frac{dL}{2w}$ vertically upwards. A similar statement can be made for the dark fringes.

4. Glass Plate*

The effect of the glass plate is to increase the optical path length of a ray from the top slit, though it essentially travels the same distance. The conditions for constructive and destructive interferences are then for the difference in OPL to be an integer multiple and half-integer multiple of λ respectively. For small angles of θ, the incident angle on the glass plate and the refracted angle are both small. The optical path length that a ray covers in the glass plate is just nt. Next, the parallel shift of a ray through the glass plate is negligible as $t \ll L$ — the rest of the path taken by a ray from the top slit is identical to that in a normal double slit experiment. Therefore, the optical path length of the top slit effectively increases by $(n-1)t$. The respective conditions for constructive and destructive interferences then become

$$d \sin\theta - (n-1)t = m\lambda,$$
$$d \sin\theta - (n-1)t = \left(m - \frac{1}{2}\right)\lambda.$$

For small values of θ,

$$\sin\theta \approx \tan\theta = \frac{y}{L},$$

where y is the vertical coordinate of a point on the screen, with respect to the center. Therefore, the locations of the bright and dark fringes for small

θ are respectively,

$$y = \frac{m\lambda L}{d} + \frac{(n-1)tL}{d},$$
$$y = \frac{\left(m - \frac{1}{2}\right)\lambda L}{d} + \frac{(n-1)tL}{d}.$$

Evidently, the glass plate shifts the fringes at small θ upwards by a distance $\frac{(n-1)tL}{d}$.

5. Wedge*

As the wedge angle is small, the distance between the two isosceles edges of the wedge is small for small vertical distances from the tip. Therefore, the effect of refraction on the path length covered inside the wedge can be neglected in this regime. As a ray impinges on the left face of the wedge, part of it is reflected while another part is transmitted. The transmitted portion then travels to the other face and part of it is reflected from that again — later interfering with the first reflected portion. This interference produces the intensity maxima and minima. The additional optical path length covered by the secondary reflected ray is $2 \times 2y \tan\frac{\alpha}{2} \approx 2y\alpha$. However, remember that the primary reflected ray undergoes a π-radian phase shift during reflection (while the secondary ray doesn't) which is equivalent to it having covered an additional optical path length of $\frac{\lambda}{2}$ where the monochromatic wavelength is defined as λ. The conditions for constructive and destructive interferences are then, respectively,

$$2y\alpha = \left(n - \frac{1}{2}\right)\lambda,$$
$$2y\alpha = n\lambda,$$

where n is the order number. Substituting $y = y_2$ when $n = 2$ into the first equation,

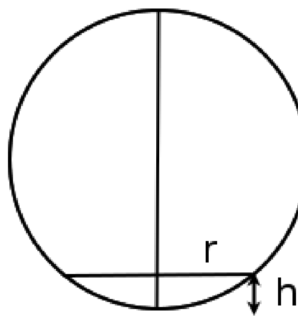
$$2y_2\alpha = \frac{3}{2}\lambda.$$

Let the third intensity maximum occur at $y = y_3'$. Then,

$$2y_3'\alpha = 3\lambda$$
$$\implies y_3' = 2y_2.$$

6. Newton's Rings**

Consider the side view of the lens (including the complete circle of radius R) and let the thickness of the air gap as a radius r from the vertex be h.



The intersecting chords theorem states that

$$h(2R - h) = r^2.$$

As the air gap is thin, h^2 is negligible as compared to r^2. Then,

$$h \approx \frac{r^2}{2R}.$$

Consider a vertical ray that impinges normally on the lens at a radius r. Part of the ray is reflected off the concave surface of the lens while some is transmitted. The transmitted portion is then reflected off the surface of the table and then transmitted back into the lens — interfering with the portion that was reflected the first time. The path difference between these portions is $2h$ (there is a negligible angle of deviation due to refraction as h is small). Therefore, the conditions for constructive and destructive interferences, while taking into account the π-phase shift due to the reflection off the table (there is no π-phase shift for the reflection off air which is optically less dense than the lens), are

$$2h = \left(n - \frac{1}{2}\right)\lambda,$$

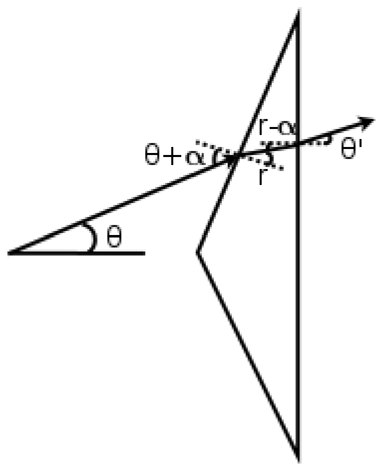
$$2h = (n - 1)\lambda,$$

respectively where n is a positive integer and corresponds to the order number. Expressing h in terms of r, while using the subscripts b and d to denote bright and dark rings, the respective radii of the nth bright and dark rings are

$$r_b = \sqrt{\left(n - \frac{1}{2}\right) R\lambda},$$

$$r_d = \sqrt{(n-1)\, R\lambda}.$$

7. Fresnel's Biprism**



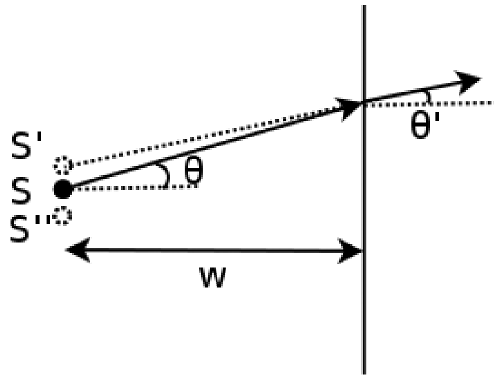
Consider a ray that is emitted from S at a small angle θ with respect to the horizontal. Its angle of incidence with the slanted surface of the biprism is $\theta + \alpha$. The angle of refraction within the biprism is given by Snell's law as

$$r = \frac{\theta + \alpha}{n},$$

where we have used the small angle approximation $\sin x \approx x$. The second angle of incidence is $r - \alpha$. Therefore, the angle that the transmitted ray subtends with the horizontal is

$$\theta' = n(r - \alpha) = \theta - (n-1)\alpha.$$

The original ray is deflected by an angle $(n-1)\alpha$. Now, consider the following diagram where the ray emitted at angle θ above the horizontal is effectively produced by a virtual source S' above S.



We know that the ray exiting from the biprism subtends an angle $\theta' = \theta - (n-1)\alpha$. Therefore, the distance between S' and S is

$$w \tan\theta - w\tan[\theta - (n-1)\alpha] \approx w(n-1)\alpha,$$

by the small angle approximation $\tan x \approx x$. Furthermore, it can be seen that all rays with small θ above the horizontal seemingly emanate from S', as the above expression is independent of θ. By symmetry, there should also be a virtual source S" below S, due to refraction with the bottom slanted surface. The two sources act as two infinitesimal slits in the double slit experiment and thus produce an interference pattern. The "slit separation" in this case is

$$d = 2w(n-1)\alpha.$$

The condition for constructive interference is

$$d\sin\theta = m\lambda.$$

For small θ, the mth-order bright fringe corresponds to an angle

$$\theta = \frac{m\lambda}{d} = \frac{m\lambda}{2w(n-1)\alpha}.$$

Each virtual source has a limited angle of illumination. S' is produced due to refraction with the top slanted side of the biprism. Its rays cover all positive $\theta' \le \frac{\pi}{2} + (1-n)\alpha$ but only a small region for negative θ'. In the boundary case, a ray emitted at an infinitesimal positive angle above the horizontal leaves the biprism at an angle $(n-1)\alpha$ below the horizontal. Therefore, the interference region corresponds to $|\theta| \le (n-1)\alpha$ (this inequality justifies the small angle assumption too). Since $\theta = \frac{m\lambda}{2w(n-1)\alpha}$, the maximum m is then

$$m = \left\lfloor \frac{2w(n-1)^2\alpha^2}{\lambda} \right\rfloor.$$

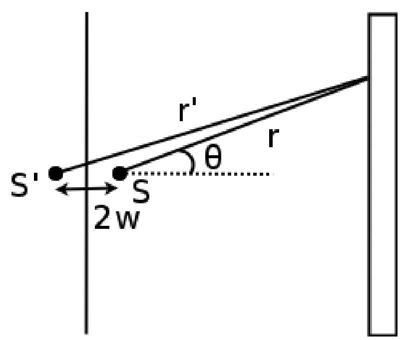
The total number of bright fringes is then

$$2\left\lfloor \frac{2w(n-1)^2\alpha^2}{\lambda} \right\rfloor + 1,$$

(plus one for the zeroth-order maximum).

8. Mirror Interference**

The mirror image of S is a virtual source S' depicted in the diagram below. The rays produced by S and S' overlap and interfere on the screen.



Consider a point on the screen that is at an angle θ with respect to S. The path length travelled by a wave from S is

$$r = \frac{l}{\cos\theta},$$

while that from S' is given by the cosine rule (discarding the second-order term $\frac{w^2}{l^2}$ en route) as

$$r' = \sqrt{r^2 + 4w^2 + 4rw\cos\theta} = \frac{l}{\cos\theta}\sqrt{1 + \frac{4w\cos^2\theta}{l}} \approx \frac{l}{\cos\theta} + 2wcos\theta.$$

The path length difference is thus

$$\delta = 2w\cos\theta.$$

The condition for constructive interference is

$$\delta + \frac{\lambda}{2} = n\lambda,$$

where the additional half wavelength takes into account the π-radian phase shift due to reflection off the mirror. Substituting the expression for δ in terms of $\cos\theta$ and $w = k\lambda$,

$$\cos\theta = \frac{n}{2k} - \frac{1}{4k}.$$

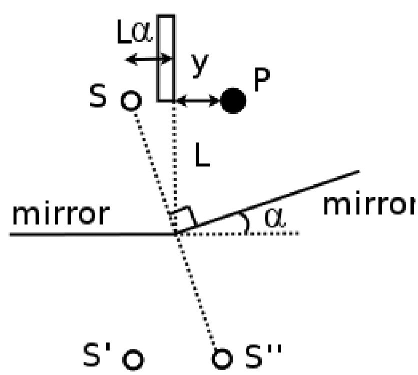
Since $0 \leq |\cos\theta| \leq 1$, the maximum order n satisfies

$$0 \leq \frac{n}{2k} - \frac{1}{4k} \leq 1$$
$$\frac{1}{2} \leq n \leq 2k + \frac{1}{2}.$$

Therefore, the total number of bright rings is $2k$.

9. Slanted Mirrors**

The two mirrors produce two mirror images S' and S" each, as shown in the figure below.



The horizontal distance between S and the center line is

$$L\tan\alpha \approx L\alpha$$

for small α. Therefore, the distance between the two virtual sources is $2L\alpha$. Furthermore, they are in phase as the reflected rays from the corresponding mirrors both undergo a π-radian phase shift due to hard reflection. The two virtual sources S' and S" then act as a double slit source with slit separation $d = 2L\alpha$ that impinge on a screen a distance $2L$ away. For constructive interference,

$$d \cdot \sin\theta = n\lambda,$$

where θ is the angle subtended by a ray from either virtual source to point P (note that the far field approximation holds as $2L \gg 2L\alpha$). Applying the small angle approximation,

$$\sin\theta \approx \tan\theta = \frac{y}{2L}.$$

Then,

$$\alpha y = n\lambda \qquad (n \in \mathbb{Z}^+)$$

is the condition for constructive interference (note that $n = 0$ is invalid as the ray from S' is blocked). Now, notice that the angular range of S' is limited — the maximum y that its rays can reach is $L\alpha$ and this occurs when the ray from S hits the intersection of the two mirrors. Therefore, the total number of locations where constructive interference occurs is

$$\left\lfloor \frac{L\alpha^2}{\lambda} \right\rfloor.$$

10. Sagnac Interferometer**

For the clockwise cycle of light, the velocities of the mirrors due to rotation, along the edges of the square, are aligned with the light ray. For the anti-clockwise cycle, the velocities of the mirrors oppose that of the light ray. The velocities of the mirrors along the edges are $\frac{\sqrt{2}}{2} \cdot \omega L \cdot \frac{\sqrt{2}}{2} = \frac{\omega L}{2}$. Therefore, the time taken for the clockwise beam to be recombined at the splitter is

$$t_1 = \frac{4L}{c - \frac{\omega L}{2}}.$$

The time taken by the anti-clockwise beam is

$$t_2 = \frac{4L}{c + \frac{\omega L}{2}}.$$

This difference in time taken and thus difference in distance traveled by the beams leads to a phase difference. Constructive interference occurs when the time difference is a multiple of the period. As L is the smallest length for which constructive interference happens,

$$t_1 - t_2 = \frac{4L}{c - \frac{\omega L}{2}} - \frac{4L}{c + \frac{\omega L}{2}} = T.$$

Performing a binomial expansion for the middle expression,

$$\frac{4L}{c\left(1-\frac{\omega L}{2c}\right)}-\frac{4L}{c\left(1+\frac{\omega L}{2c}\right)}\approx\frac{4L}{c}\left(1+\frac{\omega L}{2c}-1+\frac{\omega L}{2c}\right)=\frac{4\omega L^2}{c^2}.$$

Thus,

$$\frac{4\omega L^2}{c^2}=T$$

$$\omega=\frac{c^2T}{4L^2}.$$

11. Accelerating Cars**

Let the waves that destructively interfere (as they are traveling in opposite directions and have zero phase difference) and are received by the observer at t' be emitted by the cars with negative and positive x-coordinates at times t_1 and t_2 respectively. The total distances traveled by the emitted waves, before they reach the observer, are $c(t'-t_1)$ and $c(t'-t_2)$. Now, to determine the phase shift due to one wave, we simply have to divide the corresponding path length by the wavelength and multiply it by 2π. However, the wavelength must be Doppler-shifted. The speed of a car at time t is

$$v(t)=v_0-at.$$

Therefore, the wavelength of an emitted wave at time t is

$$\lambda(t)=cT-vT=(c-v_0+at)T,$$

where cT is the wavelength if the sources were not moving and vT is the distance covered by the sources in a single period. Note that the wavelength can be defined only because the frequency of the source is high such that waves can be envisaged to be emitted at every instance. The phase difference between the two waves emitted at t_1 and t_2 by the left and right cars when they reach the observer at $x=d$ is then

$$\frac{2\pi c(t'-t_1)}{\lambda(t_1)}-\frac{2\pi c(t'-t_2)}{\lambda(t_2)}=\frac{2\pi c(t'-t_1)}{(c-v_0+at_1)T}-\frac{2\pi c(t'-t_2)}{(c-v_0+at_2)T}$$
$$=\frac{2\pi c(t_2-t_1)[at'+(c-v_0)]}{(c-v_0+at_1)(c-v_0+at_2)T}.$$

Equating the numerator to zero and noting that $t_1\neq t_2$ (waves emitted at the same time would not reach the observer simultaneously),

$$t'=\frac{v_0-c}{a}.$$

Now, let us adopt another perspective to the problem. We can express t_1 and t_2 in terms of t'. For t_1, the distance covered by the left car from time 0 to time t_1 plus the distance covered by the wave from t_1 to t' should be equal to $l + d$. Thus,

$$v_0 t_1 - \frac{1}{2} a t_1^2 + c(t' - t_1) = l + d.$$

Solving and substituting $t' = \frac{v_0 - c}{a}$,

$$t_1 = \frac{v_0 - c + \sqrt{(c - v_0)^2 - 2a\left(l + d - \frac{c(v_0 - c)}{a}\right)}}{a},$$

where the other solution is infeasible when $l + d - \frac{c(v_0-c)}{a} < 0$ or $v_0 > \frac{a(l+d)}{c} + c$ such that the expression in the square root is larger than $v_0 - c$ (causing the other root to be negative). Similarly, the solution

$$t_2 = \frac{v_0 - c + \sqrt{(c - v_0)^2 - 2a\left(l - d - \frac{c(v_0 - c)}{a}\right)}}{a}$$

uniquely exists when $v_0 > \frac{a(l-d)}{c} + c$. The overall condition for the observer to only receive one wave from each source at time t' is hence $v_0 > \frac{a(l+d)}{c} + c$.

Bibliography

[1] Ricardo B. Comprehensive analysis of the failure of intuition in elementary rigid body dynamics. *Applied Physics Research*, 8(1), 2016.

[2] Ricardo B. Using scaling to compute moments of inertia of symmetric objects. *Eur. J. Phys.*, 36(5):055003, 2015.

[3] Ricardo B. and Lee P. Maximizing kinetic energy transfer in one-dimensional many-body collisions. *Eur. J. Phys.*, 36(2):025013, 2015.

This page intentionally left blank

Index